002: Aesthetics of New AI Interfaces

February 2021

002: Aesthetics of New AI Interfaces

This reader pulls together published writing by (or chosen by) panelists who join the Creative AI Lab's event on February 12 2021, *Aesthetics of New AI Interfaces,* a discussion held by Serpentine. Please note that texts appear as they were originally published.

Please also see panelist, Rebecca Fiebrink's free, open-source software for real-time, interactive machine learning at wekinator.org

Creative AI Lab

The Creative AI Lab is a collaboration between Serpentine R&D Platform and the Department of Digital Humanities, King's College London. Our research currently investigates: AI tools supporting artistic practices; The changing nature of artistic and curatorial practices as a result of working with AI/ML; Creative AI as a critical practice; Aesthetics of AI/ML. The Lab is funded in part by the Arts and Humanities Research Council.

Dr Mercedes Bunz, Department of Digital Humanities, King's College London (Principal Investigator)
mercedes.bunz@kcl.ac.uk

Eva Jäger, Arts Technologies, Serpentine (Co-Investigator)
evaj@serpentinegalleries.org

creative-ai.org

Interface Mythologies
Xanadu Unraveled

Christian Ulrik Andersen

Associate Professor in Digital Design at Aarhus University working with aesthetics, software, and interface criticism. He is co-author of *The Metainterface,* MIT Press 2018, with **Søren Pold**

Andersen, Christian Ulrik and Pold, Søren (2018).
"Interface Mythologies–Xanadu Unraveled." In
Interface Critique Journal Vol.1. Eds. Florian Hadler,
Alice Soiné, Daniel Irrgang

TECHNO MYTHS

Data mining, machine learning and other disciplines involved in finding patterns of data promise a future with new insights that will enable a new mode of intelligence. However, as with much other technological marketing, this is also a myth. In our interface criticism, we propose to engage with ubiquity, openness, participation and other aspects of this intelligence as mythological constructions which are presented to us via interfaces.

Following on from Roland Barthes' seminal studies of visual culture, where he discusses everything from striptease to washing powder, we intend to engage with the illusions of technologies. In many ways it is, for instance, an illusion to believe that a computer system can really forecast everything. As with weather forecasts, predictions of traffic, browsing, and other behaviours are faulty. Machine learning works by approximation and by generating generalized functions of behaviour, which are only generalizations after all; and similarly, the data we produce is captured by technologies that constantly have to deal with the noise of many simultaneous and ambiguous actions. However, from the perspective of a mythology, the important aspect is not whether the generated algorithms work or not, but how they become part of our reality. For instance, they function as speech acts that create correlations between 'data analytics' and 'intelligence', and this performative act may have a real impact when we rely on this alleged intelligence – when we market products, control traffic, fight terrorism or predict climate changes.

The mythologization of technology that takes place in the speech acts does not imply that how the technology 'really works' is hidden, but merely the ability to automatically associate certain images with certain signification in an absolute manner. To follow on from Roland Barthes, the mythologization of our smart technologies removes the history of intelligent systems, smartness, ubiquitousness, openness, and so forth, from the linguistic act. Just as we do not question that Einstein's famous equation, and equations more generally, are keys to knowledge – as Barthes describes – intelligent systems for smart cities, state security, logistics, and so on suddenly appear absolute.[1] Along with openness, participation and other techno myths, 'smartness' appears as an algorithmic reality we cannot question.

However, all techno myths should be seen as expressions of how we want the world to be, rather than what it really is. In order to perform an interface criticism, we do not need to discuss if the technologies are true or false – for the smart techniques of data mining, machine learning, and so forth, obviously work – but we need to realize that their myths are also part of our reality. As Philip Agre

1 Roland Barthes, *Mythologies*, transl. Annette Lavers (New York: Hill and Wang, a division of Farrar, Straus & Giroux, 1972).

has noted, we subject our actions to the system that needs to capture them as data; and this deeply affects the way we produce, socialize, participate, engage, and so on.[2] The monitoring of academic production and the capture of citations is, for instance, used to create indexes which indicate impact. Ideally, this can affect the efficiency of academia and be a relevant parameter for funding opportunities, careers, and the like. Even though this efficiency may be absent, the data capture still has an effect on the perception and performance of academic work; it is constitutive of our habitat and subtly affects our habits.

In many ways, the technological myths always feel real, and are dominant actors that affect a range of areas – from the perception of the weather, to our cities, and our cultural production and consumption. We have every reason to question not only if the technology works, but also the implications of its myths. It is often when we realize the pointlessness of our actions (that texts can be quoted for their mistakes, rather than their insights; or their summaries of knowledge rather than their epochal value) that we structurally begin to question the absolute assertions about the world embedded in the myth, and also to envision alternatives.

In this article, we do not want to dismiss intelligent, open, participatory or other technologies, but to discuss how technologies participate in the construction of myths. To us, this criticism fundamentally involves a mythology – a critical perspective on the interface that explores how the interface performs as a form of algorithmic writing technology that supposedly transcends signs, culture and ideology. To focus on the interface as a a language diverts attention away from technology's immediate assertions about reality – the technical fix – and highlights the materiality of their staging. The aim will be to discuss how technologies perform as dreams of emancipatory or other post-semiotic idealized futures, and argue for the need for an interface mythology that critically addresses the technologies as myths; and unravels them as value systems and tools for writing – of both future functionalities and future cultures.

DREAM MACHINES

There is a general tendency to develop technology in the light of cultural utopias. The development of hypertext is a very good example of this. With the emergence of hypertext in the sixties (and later the WWW, weblogs, social media, and much

---

2 Philip E. Agre, "Surveillance and Capture: Two Models of Privacy," in *The New Media Reader*, ed. Noah Wardrip-Fruin and Nick Montfort (Cambridge, Massachusetts and London, England: MIT Press, 2003). According to Agre there are two dominant notions of surveillance. Surveillance is often perceived in visual metaphors (i.e., 'Big Brother is watching'); however, computer science mostly builds on a tradition of capturing data in real time, and is often perceived in linguistic metaphors ('association', 'correlation', etc.). Hence these metaphors are also better suited to describe the kinds of surveillance taking place when data capture permeates social life, friendship, creative production, logistics, and other areas of life.

more), the development of various forms of textual networks has been intrinsically linked to strong visions of new ways of producing, experiencing and sharing text. One of the strongest proponents of such visions has been Theodor H. (Ted) Nelson. Nelson's *Xanadu* is a lifelong project, and it has been the outset for numerous reflections on the development of hypertext. Perhaps the most well-known of these texts is *Computer Lib/Dream Machines* from 1974, a self-published book featuring illustrations, cartoons and essays on various topics, all aiming in different ways to explore alternative ways of thinking related to computers.

Furthermore, the book can be read from both ends. The one end offers a technical explanation for common people of how computers work; as Nelson writes: "Any nitwit can understand computers, and many do. Unfortunately, due to ridiculous historical circumstances, computers have been a mystery to most of the world."[3] The other end is meant to make the reader see the development of the computer as a "choice of dreams."[4] According to Nelson, what prevents us from dreaming is the developer's incomprehensible language (or, as he labels it, "cybercrud"), which in his view is just an excuse to make people do things in a particular way; that is, to let the technocratic visions of culture stand unchallenged.

Already in 1965 Nelson invented the term hypertext for a new kind of file structure for cultural and personal use:

> The kinds of file structures required if we are to sue the computer for personal files and as an adjunct to creativity are wholly different in character from those customary in business and scientific data processing. They need to provide the capacity for intricate and idiosyncratic arrangements, total modifiability, undecided alternatives, and thorough internal documentation. [...] My intent was not merely to computerize these tasks but to think out (and eventually program) the dream file: the file system that would have every feature a novelist or absentminded professor could want...[5]

In this way, Nelson was already in 1965 aware that developing alternative uses of the computer was closely linked to developing alternative versions of the technical structure and even the file system. He continued – and still continues – to develop his idea of hypertext, of which he premiered the first publicly accessible version at the *Software* exhibition of technological and conceptual art in New York in 1970. Visions and dreams appear in a recognition that the power of computation – or of computer liberation – is linked to visions of a new medium; that the inner

3 Theodor H. Nelson, "Computer Lib / Dream Machines," in *The New Media Reader*, ed. Nick Montfort and Noah Wardrip-Fruin (Cambridge, MA: MIT Press, 2003 (1974/1987)), 302.

4 Ibid. 305.

5 "A File Structure for the Complex, the Changing, and the Indeterminate," in *The New Media Reader*, ed. Nick Montfort and Noah Wardrip-Fruin (Cambridge, MA: MIT Press, 2003 (1965)), 134.

signals of cathode ray tubes are related to signs and signification, and therefore to cultural visions. In other words, they are linked to the hypothesis that the computer interface, at all levels, and not just the graphical user interface, is an interface between the technical and the cultural. When text, for instance, is treated by protocols there is a double effect, where not only the cultural form of the text changes (e.g. from book to hypertext), but also the technology itself appears as a deposition of cultural values. This is why the discussion of the future of text and images, on the web and in e-books, also appears as a discussion of text protocols and formats.

THE SUBSUMPTION OF DREAMS

Many writers and theorists have adopted Nelson's visions of alternatives, and of new modes of producing, reading and sharing text. For example, in his book *Writing Space*, Jay Bolter explored what writing was before and potentially could be with hypertext.[6] Bolter's main hypothesis was that print text no longer would decide the presentation and organisation of text, and that it no longer would decide the production of knowledge. Readers would become writers, and this would undermine the authority of print text; writing would become liquid, and we would experience a space of creative and collective freedom. However, as we have experienced on today's Internet, not everything seems as rosy. There are plenty of reasons to look more critically at Facebook, Twitter, Wikis and other services.

Nelson's Xanadu system had already included an advanced management instrument, the so-called 'silver stands': stations where users can open accounts, dial up and access the information of the system, process publications and handle micro payments. Nelson himself compares this to a McDonald's franchise and the Silver Stands somehow resemble the Internet Cafés of the late 90s and early 2000s or the commercial, centralized platforms of Web 2.0. Furthermore, copying content in the Xanadu system is restricted to dynamic "transclusions" that include the current version of the original text and assure a small royalty when accessed, a so-called "transcopyright".

When looking at the services of Facebook, Google, Amazon, Apple, and so on today, it is similarly obvious that the common production modes characteristic of a free writing space are accompanied by strict control mechanisms. There are, for instance, strict protocols for the sharing, searching, writing and reading of text, and these protocols often ensure an accumulation of capital and compromise the anonymity and freedom of the participant. In other words, the in-

6 J. David Bolter, *Writing Space the Computer, Hypertext, and the History of Writing* (Hillsdale, N.J: L. Erlbaum Associates, 1991).

strumentalization of the dream includes everything else but the dream. The envisioned shared, distributed, free and anonymous writing space is in fact a capitalised and monitored client-server relation.

This critique of contemporary interface culture is perhaps not news, but what we want to stress here is the effect of the instrumentalization of dreams and visions. What this indicates is that down the 'reactionary path' (that is, the path of instrumentalization), our dreams turn into myths. However, the ethos of the dreams remains, and become automatically associated with the technical systems.

THE THREE PHASES OF MEDIA TECHNOLOGIES

The dream of a shared writing space, a Xanadu, that overcomes the problems of representation facing linear text forms, as well as the hypertext system's instrumentalization of this dream, the mythological status of such systems, and the adherent critique of them, all fit into a three-phase model of media presented by the German media theorist Harmut Winkler.

From a linguistic perspective all new media are, in the first phase, considered post-symbolic, concrete and iconic communication systems that present a solution to the problem of representation, or the arbitrariness of the sign. Winkler even sees the development of media as "deeply rooted in a repulsion against arbitrariness", and a "long line of attempts to find a technical solution to the arbitrariness" dating back to the visual technical media of the 19th century.[7] In addition, hypertext was perceived as establishing a more true relation between form and content, because of its more intuitive, democratic, and less hierarchical, nonlinear structure. It will often be the investment in the dreams that pays for their technical implementation: You not only buy new functionality, you buy a new way of living, working, thinking and dreaming. In this way, the development of hypertext, the WWW, social media – and also computer games and virtual reality, and their alleged liberation of the user – is driven by an urge to fulfil a dream, a vision of a new future.

In the second phase, the utopias become natural, stable and hegemonic. Through subsumption by market forces they become commodified, and sold as myths of being part of a media revolution. However, the subscription to this reality also contains an explicit lack of visions of alternative futures, and is therefore also without the critical, activist and heroic dimensions of the first phase.

It is, however, also a phase where people begin to study the media and learn how to read and write with them. In other words, the new media begins to enter a phase where you see it as a language, and hence where the arbitrariness of the sign is

7 Hartmut Winkler, *Docuverse* (Regensburg: Boer, 1997), 214.

reinstalled. In the third phase, this arbitrariness has turned into disillusion over the media's lack of abilities; which, however, also constitutes the ground for new visions, new media technologies, new interfaces, and new media revolutions.

The question is how far are we, today, from Ted Nelson's critique of centralised data processing and IBM-like visions of efficiency and intelligence? In several ways, it seems as if we are in a phase where we might soon begin to regard big data, smart systems, social intelligence, and so forth, as a language; where we begin to see through the technological systems' mythological statuses, or at least their dark sides in the form of control and surveillance. This is by no means an easy phase. As Ted Nelson also noted, "Most people don't dream of what's going to hit the fan. And computer and electronics people are like generals preparing for the last war."[8] The developers of technology and their supporters will often insist that their system is the future, and that the users' actions need to follow the system's intrinsic logic.

INTERFACE MYTHOLOGIES

From a design perspective, the assumption will typically be that the clearer the representation of the computer signal-processes appears (or the mapping of mental and symbolic labour – the formalization of labour to computer language performed by the programmer), the more user-friendly and understandable the user interface appears. To computer semiotics, the aim was ultimately to create better interface design. However, in relation to an interface criticism, it is noteworthy how computer semiotics also explains how a design process in itself contributes to the mythological status of the interface – its absolute assertions about the world.[9] In other words, the myths of interfaces are not only established through how they are represented elsewhere (how they are talked about, written about, advertised, etc.), but also through the interfaces themselves, and how they are designed. It is in its design as a medium, and in its claims of an iconic status as a communication system, that we find the interface's operationalized mythology. And, in a general perspective, this is not unlike how media such as photography, film, the panorama, and so on, according to Harmut Winkler, have tried to operate in earlier times.

To read this myth demands that one begins to read the media – or, in our case, the interface. It is a tool for reading and writing, and not an absolute representation of the world. We must, therefore, begin to pay attention to the establishment of sign-signal relations that take place in the interface design, as a particular production mode, a particular kind of labour; a production of signs that at

8 Nelson, "Computer Lib / Dream Machines," 305.

9 On computer semiotics and the work of Frieder Nake and Peter Bøgh Andersen, see Søren Pold and Christian Ulrik Andersen, *The Metainterface: The Art of Platforms, Cities and Clouds* (Cambridge, MA and London, England: MIT Press, 2018).

once reflects cultural and historical processes, and leaves an imprint on the world and how we organise and deal with it.

For instance, the software of the print industry, as Nelson also demonstrates, both reflects the historical and cultural origins of print and negotiates the reality of text, as searchable, sequential, iterative, sortable, and so forth. Our file formats and standards for storing and showing data also reflect such processes. Jonathan Sterne, for instance, has recently analysed how the diameter of the Compact Disc directly reflects relations to the cassette tape, and how the mp3 format also holds an audio culture of listening that is embedded in the sound compression, and how this directly challenges the cnception of technological progress as equal to increased high fidelity.[10] Even the electrical circuits and the signal processes deep inside the computer can be viewed as the result of language acts, as Wendy Chun has pointed out.[11]

Computer software and its formats and platforms promise us dreams of the future, of technological progression, better opportunities to make our music portable and shareable, better ways of organising our work, and so forth. It is often these dreams that carry the technological development. However, the dreams have a tendency to freeze, and gain an air of absoluteness, and of hegemony. This happens through their commodification and appropriation to a reality of power and control. Technology is marketed as a utopia of being in the midst of a media revolution. But in this phase the cultural and historical residues are hidden. We are seduced by the interface into neglecting the work behind it, and the operationalization and instrumentalization of dreams that takes place. The interface appears mythical, absolute and frozen. We do not see the mp3 format's compression of sound as a result of an audio culture, but as the only possible scenario, a technological fact; and we do not see the IT systems of workers as the result of a negotiation of labour processes, and we do not see the operational system's metaphorization of actions as other than a result of natural selection in the evolution of technologies. To get out of the deception of the technological facts we need interface mythologies – critical readings of the interface myths.

---

10 Jonathan Sterne, *Mp3: The Meaning of a Format, Sign, Storage, Transmission* (Durham: Duke University Press, 2012).

11 Wendy Hui Kyong Chun, *Programmed Visions: Software and Memory* (Cambridge, MA and London, England: MIT Press, 2011).

REFERENCES

Agre, Philip E. "Surveillance and Capture: Two Models of Privacy." In: *The New Media Reader*, edited by Noah Wardrip-Fruin & Nick Montfort, 737-60. Cambridge, MA: MIT Press, 2003.

Barthes, Roland. *Mythologies*. Transl. Annette Lavers. New York: Hill and Wang, a division of Farrar, Straus & Giroux, 1972.

Bolter, J. David. *Writing Space the Computer, Hypertext, and the History of Writing*. Hillsdale, N.J: L. Erlbaum Associates, 1991.

Chun, Wendy Hui Kyong. *Programmed Visions: Software and Memory*. Software Studies. Cambridge, MA: MIT Press, 2011.

Nelson, Theodor H. "Computer Lib / Dream Machines." In: *The New Media Reader*, edited by Nick Montfort, and Noah Wardrip-Fruin, 301-38. Cambridge, MA: MIT Press, 2003 (1974/1987).

Nelson, Theodor H. "A File Structure for the Complex, the Changing, and the Indeterminate." In: *The New Media Reader*, edited by Nick Montfort and Noah Wardrip-Fruin, 133-45. Cambridge, MA: MIT Press, 2003 (1965).

Pold, Søren, and Christian Ulrik Andersen. *The Metainterface: The Art of Platforms, Cities and Clouds*. Cambridge, MA: MIT Press, 2018.

Sterne, Jonathan. *Mp3: The Meaning of a Format (Sign, Storage, Transmission)*. Durham: Duke University Press, 2012.

Winkler, Hartmut. *Docuverse*. Regensburg: Boer, 1997.

The Force of Communication

Mercedes Bunz

Creative AI Lab Principal Investigator and Senior Lecturer in Digital Society at the Department of Digital Humanities, King's College London. Her research explores how digital technology transforms knowledge and power.

Bunz, M. "The Force of Communication",
in: P. Bialski, F. Brunton, M. Bunz (eds) Communication. University of Minnesota Press 2019: 51-91.

Abstract This chapter examines our contemporary dialogues with digital media. Starting with the observation that a certain force has always been a theme in theories of communication, it explores what kind of force unfolds when we communicate using digital interfaces. Looking at the communication of digital technology through the perspective of Althusser's theory of interpellation, it becomes apparent that at the beginning of the 21st century digital interfaces address us as very young children. Exploring this phenomenon of infantilization further, the chapter turns to the history of graphical user interfaces and to the influence of child psychologist Jean Piaget on computer scientists, especially on Seymour Papert and Alan Kay. It can then be shown that the force of this particular way of addressing leads to the paradox of two very different effects that can appear simultaneiously: Being addressed as a very young child can patronize as much as it can invite the user to a playful learning thereby enhancing knowledge about those interfaces. The final part of this chapter explores this paradox theoretically.

The things around us, having become media, have started to address us. Their first utterances went unnoticed: for years, our cars have loudly insisted that we fasten our seat belts. Informed by sensors, they scream as if they feared for their bodies while being parked or shout for help when they reckon that someone else, whom they do not know, wants to take them. This mode of communication quickly spread to the house. Now the robotic vacuum cleaner eagerly informs us when it is stuck and asks us to "move Roomba to a new location." And driven by new advances in natural language processing I have explored elsewhere (Bunz and Meikle 2018, 45–67), intelligent personal assistants with names like Siri and Alexa wake up to address us when they hear someone calling their names—in contrast to our fellow humans, who ignore everyone around them while under the spell of a screen. When things became interactive, they established a new kind of dialogue with us, the humans. To use technical interfaces today means to communicate with technology. Of course, it is not technology itself that has raised its head and started to speak. Even though it has learned to communicate, it has not become a human subject, although it has always been more than an object. Heidegger ([1954] 1977, 4) had good reason to look *further* into the agency of technology by reconsidering what is usually taken for granted—“technology is a means to an end. . . . Technology is a human activity”—thereby questioning the instrumental definition of technology. Now that our technological devices have started to address us with multiple voices, we need to continue his analysis. So in what way can we investigate how technology addresses us without thinking it is speaking to us? For this is certain: when technology starts to speak, it is not technology we hear. Still, this is a development that is transforming our contemporary discourse and, with it, what can be called our “being with technology.” This essay explores the force of digital communication, starting with a methodological discussion of how to approach technology. Having clarified this, it then links different aspects together: communication theories and the way we are addressed by digital media, child psychology and computer science, interface design and political theory. But let us start this endeavor by looking at what happens—what forces speak—when we communicate.

Being with Technology

Communication theories have always suspected that communicating with media *transforms our being in this world* in various ways. This section approaches these theories and this transformation in three ways. First, it summarizes historical theories of communication to foreground their common assumption, namely, that there is a *force* happening when we communicate. To understand where this force is generally located when it comes to digital technology, it then turns to contemporary theories. Finally, it discusses technology as a situation: the situation of being addressed by digital technology. But let’s start with historic takes on communication.

Over the years, theorists have developed very different takes on communication. Yet, one assumption has always been at the heart of all theories: there is a force happening while we communicate. The following communication theories illustrate this, although the list is by no means exhaustive:

Shannon. An interest in the force of communication can already be noticed in one of the early theoretical takes on communication, in Claude Shannon and Warren Weaver's (1949) *The Mathematical Theory of Communication,* which my coauthor Finn Brunton discusses with brilliance and in more detail in chapter 1. Their theoretical concept of information implies that the capacity of a medium defines its possibilities to produce meaning, thereby claiming a certain dependency on the transmitting medium. Inspired by their theory, the German media theorist Friedrich Kittler (1999, xxxix) would condense this later to the claim that "media determine our situation, which—in spite or because of it—deserves a description."

Derrida. The French philosopher adds to this perspective (that something else is going on when communication is happening) by observing that communication also does not simply transmit content. As he points out in his well-known essay "Signature Event Context" (Derrida 1977), sending a message relies on its fundamental capacity for displacement. The fact that a message functions after it has been sent from A to B means that it "breaks with its context" (9) and has an "iterative structure, cut off from all absolute responsibility." In other words, one can never be certain of its meaning.

Williams. The cultural critique points again to a very different aspect, one more related to the link of communication with "communion." In his *Keywords: A Vocabulary of Culture and Society,* Williams (1985, 72) discusses the force of communication that lies in its distributive act: "make common to many, impart." When communication makes something common to many, however, two very different things can happen, as Williams points out: it can "transmit" in "a one way process" or "share" (72). In this capacity, communication has the force to manipulate as well as to integrate and foster participation.

Haraway. Not far from this position, we find the important take of Donna Haraway on communication technologies. In “A Cyborg Manifesto” (Haraway 1991), she points to a very specific force by showing that communication technologies create social relations that structure our identity, which means that they can also restructure it. Haraway thus points out that they can be “crucial tools recrafting our bodies” and that “they should also be viewed as instruments for enforcing meanings” (Haraway 1991, 164). According to her, communication can be a discursive weapon.

Although the preceding approaches articulate very *different* perspectives and motives, all of them notice *a force* happening when there is communication—a force that is shaping our situation through shaping the possibilities of communication (Shannon and Weaver 1949), a force that can never be fully controlled (Derrida 1977) and, from a very different perspective, a force that can reach but also manipulate the many (Williams 1985) as much as it can be used as a weapon (Haraway 1991) to restructure our discourse. This chapter continues their productive suspicion that communication is always more than a transparent exchange of information. By looking into the specific case of digital technology, it explores the hypothesis that the rise of digital media is accompanied by a specific force, which differentiates it from other technologies. To enquire about this, it is necessary first to look into the theoretical setup of digital media. Can such a force also be located when it comes to digital technology?

When approaching this question, one quickly notices a rather confusing situation. Recent studies of digital technology (Bratton 2016; Chun 2016; Crawford and Joler 2018; Gitelman 2013; Starosielski 2015) have rightly pointed out a feature specific to digital communication, which is shaped by a situation far more complex than a "communication channel." Bratton (2016) has most explicitly developed this thought, showing that the technical layers of the internet's OSI architecture, by now grown into a network of planetary scale, can be described as a "stack." To explore communication, different layers of this "stack" must be taken into account: the material communication layer providing energy and matter, controlled by an optimization layer and used by an application layer (53), for example. Here network communication challenges previous theories of software.

Being written in code, software has been organized by two strands of communication and, with it, two interfaces: one for the machine (an interface whose alienness Finn Brunton explores in chapter 1 of this volume) and one for the user (an interface whose alienness I explore here). Their conflating layers are the reason why Wendy Chun (2011, 3), informed by her double degree in both systems design engineering and English literature, has called software "a notoriously difficult concept":

> Software perpetuates certain notions. . . . It does so by mimicking both ideology *and* ideology critique, by conflating executable with execution, program with process, order with action. Software, through programming languages that stem from a gendered system of command and control, disciplines its programmers and users, creating an invisible system of visibility. (Chun 2008, 316)

The disciplinary machine that software is affects programmers and users alike, as Chun points out. Following her, Alexander Galloway (2012) has addressed the interface as effect and ethos to make a similar point: interfaces do not simply transmit our messages; instead, they open—or enforce?—a very particular dialogue with technology, a point that needs to be pondered for a moment.

When discussing digital media, media theorists have often differed over where the force of digital technology originates. That there is a force, they agree—the algorithmic, as, for example, Rita Raley (2016) pointed out in her precise essay on algorithmic translation, is not purely mechanical. But where is it that media and technology scholars have to look? Do they need to look at the code with which a programmer is communicating and to which Paula Bialski turns in chapter 3? Or is the force located in the graphical user interface communicating with the user? When approaching digital technology, we too often follow "the logic of what lies beneath," as Chun (2011, 20) notes, even though "code is also not always the source, because hardware does not need software to 'do something'" (25). To make things even more complicated, further technological developments have stressed different parameters, such as data (Gitelman 2013) or machine learning architectures (Mackenzie 2017), and more parameters at the moment still unknown will follow. Thus, when looking at digital technology, this chapter assumes that for the process of communication, multiple interconnected layers are playing a part. Being interested in a very specific aspect of our dialogue with technology, however, this chapter does not focus on each of those layers but studies one particular moment: the moment when technology is addressing us. Whereas Brunton before me turns to Licklider to explore the complex setup that enables machines to communicate with each other, and Bialski in the next chapter turns to programmers to study the code review process, my chapter looks at the situation that enfolds when machines communicate with us. For this, it first needs to clarify its method of approaching technology.

As stated earlier, when technology communicates with us, it is not technology itself that raises its head and starts to speak—technology is *not an acting subject.* As Heidegger has pointed out, technology has also always been *more than an object*; that is, it has always been more than a means to an end. If it is neither a subject nor an object, however, how can in our case the force of communication regarding digital technology be approached? Here Hannah Arendt's ([1958] 1998, 151) short take on the problem of technology, which she develops while discussing the transformation of human life through technology, points our thoughts in an interesting direction:

> The discussion of the whole problem of technology, that is, of the transformation of life and world through the introduction of the machine, has been strangely led astray through an all-too-exclusive concentration upon the service or disservice the machines render to men. The assumption here is that every tool and implement is primarily designed to make human life easier and human labor less painful. Their instrumentality is understood exclusively in this anthropocentric sense. But the instrumentality of tools and implements is *much more closely related to the object it is designed to produce.* (emphasis added)

Here Arendt states that any given technology is more closely related to another technology than to a human subject. To her, technology is driven by an immanent ("closer") relation. This does not mean, however, that technology acts as a subject that masters the human. Humans play a part in the development of technology, which becomes clear in an "important assumption" added by Arendt: "that the things of the world around us should depend upon human design and be built in accordance with human standards of either utility or beauty" (152). Pleading for human standards, Arendt shifts the focus onto technology in an interesting way. She approaches it more *as a situation* and less *as a subject,* which becomes explicit in the following quotation: "The question . . . is not so much whether we are the masters or the slaves of our machines, but whether machines still serve the world and its things" (151). This chapter follows her approach when studying the force of communication by investigating how technology as a situation can be thought of in more detail. What should be examined? How does a technical situation need to be studied? To answer these questions, the chapter links Arendt's approach to Gilbert Simondon, with whom her take on technology resonates.

Like Arendt, Simondon (2017) finds our understanding of technology fundamentally flawed. Instead of emphasizing curiosity or understanding, Simondon critically remarks that our usual approaches toward technology oppose humans and machines (15). To overcome this, he rethinks this relation. In the chapter "Evolution of Technical Reality: Element, Individual, Ensemble," he describes how technical evolution is not driven by men or machine but by an "ensemble" of the two. There is no master anymore who is in control of the process of a technical development. And this shift from a master relationship to an ensemble raises a question: instead of a gifted inventor or mad genius, what drives the development of technology?

For Simondon, similar to Arendt, the answer lies in the productive relations between men and technology, which create a process of "concretisation" (Simondon 2017, 33; also Iliades 2015). He sees this, for example, in the development of X-ray tubes: regarding the Crooks tube and its later "successor," the Coolidge tube, Simondon finds the engineer William Coolidge elaborating on technical functions of the already existing Crooks tube. Coolidge "purified" them to improve the tube's functioning—a process of concretizations in which specific aspects of an already existing technology get further developed: "the functions are thus purified by their dissociation, and the corresponding structures are more distinct and richer" (36). Instead of being struck by a flash of genius, it is the "technical reality" of the Crook tube that inspires the new product. Thus it is the technical reality itself that fosters further development, although this reality needs the human to concretize: "machines can neither think nor experience [*vivre*] their mutual relation; they can only act upon one another in actuality, according to causal schemes." With this, the role of the human comes into play: "Man as witness to machines is responsible for their relation" (157).

Neither human nor technology can initiate the process of further development on its own. They need to relate to each other. With the human as an *enabling witness,* the relation of man and machine can be sketched as an *ensemble* instead of as an opposition. This puts the human in a very distinct role: the human is *not master* of machines digital or mechanic but their *interpreter.* In Simondon's (2017, 150) words, "man understands machines; for there to be a true technical ensemble man has to play a functional role between machines rather than above them" (see also Combes 2013, 57). Here the concrete technical relation of a technical object to its milieu describes an immanent development driven by "concretisations" that are nondirectional. Fascinated by constant technical change, Simondon (2012, 13) will later describe technology as characterized by an "*opening*": "technical reality lends itself remarkably well to being continued, completed, perfected, extended." Thus, in the middle of this, one finds an interesting tension: technology puts forth a situation that then needs a human to continue, complete, perfect, and extend it, in short, to turn it into reality. At the same time, technology follows its own, alien logic in what it offers to be continued, completed, perfected, and extended. We cannot predict the future of the technology we have invented. Even in the twenty-first century, in which we are facing a field as closely guarded as an economy driven by digital technology, we are never certain which technology will become the "next big thing."

Technology is a force alien to us that has now started to speak and process language. But just because it has started to process language and can now say something, we should not mistake it for a speaker. *Being with technology* instead means to approach technology as a technological ensemble, as a continuously developing situation made up of humans and technology. Thus we need to study what kind of situation unfolds when technology communicates with us as we aim to avoid treating technology as an anthropocentric subject that acts and/or speaks. Luckily, a blueprint for the power of communication that does not stem from a subject (although a subject is involved) can be found in the concept of interpellation Louis Althusser introduces when discussing the notion of ideology.

Althusser's notion of ideology evolves around an interesting shift. While he analyzes communication (or interpellation), he does not look at what is said or what can be said. Instead, Althusser (2014) focuses on the situation created when being addressed and the force of this address. In his essay "Ideology and Ideological State

Apparatuses," he analyzes the structural force happening in the moment of communication. Using the example of a policeman calling out to you on the street, he illustrates that communication situates (even appropriates) its participants by establishing a link between sender and receiver in the act of interpellation: it constitutes a subject. His description of this constitution has turned into a highly influential theory of interpellation, although it is less a "theory" than just a few paragraphs. In those paragraphs, Althusser shows that a specific social role—in his words, a "subject"—comes into being by "the practical telecommunication of hailings" (264). To illustrate how this "hailing" or "interpellation" functions in the context of ideology Althusser introduces an individual that turns around in response to a policeman shouting "Hey, you there!" (264) to "answer" that call. And in exactly that moment, so Althusser, one becomes a subject relative to the ideology of law and crime. In other words, in that moment, one experiences the social force of communication, which Althusser calls ideology: "ideology 'acts' or 'functions' in such a way that it 'recruits' subjects . . . . , or 'transforms' the individuals into subjects . . . by that very precise operation which I have called interpellation or hailing" (264).

In the twenty-first century, this operation of interpellation Althusser described, an operation that creates a situation of recruitment by establishing a link between a sender and receiver, is still continuing. Only now, it can be found in new and different forms of communication—and this is the hypothesis I would like to bring to a test in this chapter: Today, the recruiting of subjects happens when technology addresses us. By interacting with the interfaces of technology, we are situated through this communication and recruited as specific subjects. Of course, that we make a world for others to live in through our technological creations has been an aspect in philosophy of technology, which Langdon Winner (1986, 17) but also Donna Haraway (1997) and many others have addressed in much detail. This chapter adds to those explorations of politics we built into our technologies, although it will be slightly shifting the view. By approaching technology with Arendt as a situation and by trying to understand the contemporary technological ensemble (Simondon), it will not look at what is being said to us by technology. Instead, it is interested in the kind of situation that unfolds. As what kind of subject are we recruited in that situation? The next section therefore observes the communication with technology to tune into how something is being said when technology addresses us.

How Is Technology Addressing Us?

To capture how technology addresses us, this section analyzes three different examples partly drawing on earlier research (Bunz 2015): it looks at the introduction of Apple's iPad to study its early interface design, considers the brand communication of internet companies and their fondness of mascots, and, finally, turns to the Google Doodles that appear on the landing page of Google search, which one passes by when searching for other information.

On April 3, 2010, Apple's cofounder, chairman, and chief executive officer unveiled a tablet computer it introduced as "iPad." Its new product was operated via a touch screen and could play music, take photos, shoot video, and perform internet functions such as web browsing and emailing; more applications, from games to social networking, could be added. In its first fiscal year following the launch of the new product range, Apple sold 32 million iPads, with 140,000 apps being created for it by December 2011 (*Economist* 2011). One could say that with the success of the iPad, a new era in the relationship between human and computer materialized: the tablet computer showed that digital communication had left the workplace to become a commodity in our day-to-day lives. Computers had certainly entered leisure time with game consoles long before. The iPad, however, could be used for much more than just gaming. It could perform all tasks done by a personal office computer at that time, although it was not supposed for working. Its reduction to a large touch screen that weighed 680 grams made it comparable to a heavy book or magazine that could be read at home on the couch. It was its slick materiality that differentiated it from a computer as much as its specific user interface.

By that time, screens had been technically refined so that their visual interfaces no longer needed to be operated via minimal black-and-white icons. They could be replaced by touch screens with voluptuous 3D buttons more to the taste of Steve Jobs. As the former CEO of the animated film studio Pixar, he had a passion for reality imitating 3D graphics, as had Scott Forstall, the first architect of iOS, the software developed for the iPhone and iPad. Thus the early iPads had many 3D buttons and other skeuomorphic features each mimicking an original: the Notepad app had a border of stitched leather to make it look like a real notebook, the Podcasts app displayed a reel-to-reel tape deck when one pressed play, and the calendar and contacts apps looked like small books and featured a page-turn animation. Making apps and items mimic their real-world counterparts gave the iPad a stuffy look and feel. This continued in a different way Apple’s traditional appeal to nontechnical people. Right from the start, the company had established its computer as a fun-to-work-on machine by including features such as greeting users with a “happy Mac” when starting or by using symbols like the “dogcow” (indicating the setup of a page), scissors (for the cut command), or the trash can, which were created by Susan Kare for the back then still limited black-and-white screens. Now computers had entered a new, advanced, but also more serious era—at least that was the impression Apple gave with their design of the first iPad. Its look and feel communicated to the user that computers had come of age, although not for very long. Technically, all screens from phones to tablets to laptops to PCs were able to display complex, grown-up 3D interfaces. Still, a new and very different trend emerged that soon became more successful than mimetic skeuomorphism.

Surprisingly, the new trend was initiated by Apple’s rival Microsoft, which, after the iPhone’s success, had already been written off. Faced with the staggering success of Apple’s phone, Microsoft had to respond with an original and different approach: for their handheld devices, the Microsoft designers decided to focus on cards and not on buttons. Eager to avoid Apple's extensive use of skeuomorphism (Wingfield 2012), their inspirations came from the design principles of classic Swiss graphic design, which favors a minimal style, emphasizes typography, and uses a grid that can often be seen on European transportation signs. Instead of buttons, they used text placed on cards, which one could navigate laterally through scrolling canvases. Their typography-based design language came to be known as Microsoft design language. Its principles had originally been developed for Microsoft's mobile media player Zune (2006–8), before they were taken over to the Windows phone, launched in 2010. Although the device did not have the same success as the iPhone, its design would inspire others, Google among them—and Google's logo in fact exemplifies this new and different approach to user communication.

While Apple's skeuomorphic design for the iPad communicated its device as a toy-tool for grown-ups, the flat design Microsoft had initiated would go a very different way—and with it a new form of addressing the user would begin. Early on, Google would be part of this. On Wednesday, May 5, 2010, the search engine Google changed its logo for the first time in ten years and eleven months (Googleblog 2010). The new logo was less skeuomorphic and more colorful. Its three-dimensional letters in red, yellow, and blue, plus the green letter *l* based on the font Catull, lost their drop shadows. The logo had exchanged the rich details of skeuomorphism in their big typography with louder colors and simpler forms. Google's senior user experience designer Wiley explained the change on the search engine's blog as follows: "The new logo is lighter, brighter and simpler. We took the very best qualities of our design—personality and playfulness—and distilled them" (Googleblog 2010). Experts agreed. Already before the change, British graphic designer Peter Saville, known for minimal design like the radio signal cover for Joy Division's album *Unknown Pleasures,* described Google's logo in an interview not just as playful. For him, it was addressing children: “Everything about it is childlike: the colors, the typeface, even the name” (cited in Rawsthorn 2010).

The redesign intensified this further. Chris Moran, then the *Guardian*’s search engine editorial optimizer, commented on the new look and feel as a turn toward “My First Search Engine” (pers. comm., May 6, 2010). Online, the rise of flat design had begun, even though it would take a while before its triumph over skeuomorphism became recognizable—it was not until 2013 that an animated web page displayed the “battle flat design vs. realism” (Intacto 2013). Flat design opposed skeuomorphic and other “artificial” design techniques in favor of two-dimensional, “flat” illustrations; big typography; and bright colors for a more simplified aesthetic. When the new design became a mainstream trend, however, something else changed—technology would approach the user in a different way. The new design style addressed a very different user—not an adult one. Visually, the style resembled books for very young children. Addressing the user as a very young child, however, was a transformation that did not happen abruptly and not just in one field. With hindsight, years before 2013, the new trend in brand design could have been spotted on the World Wide Web. And although it went unnoticed for a long time, it fundamentally changed how brands approached the user.

Contemporary brand communication generally has a double function: it enables the user to identify a product and, for this, gives the product or service a specific identity or image (Millman 2012; Holt 2004). With the internet, as many marketing books were eager to explain (Levine et al. 2000), brands had to become a conversation. But this was not the only novelty. Online, the rules seemed to be different, which is why several internet companies embraced animals (or aliens). Or was it because they addressed *someone* very different? In any case, if one attentively observed the brand communication of “online” products and services, one could notice that animals had peacefully appeared in large numbers. Next to the fox of the web browser Firefox chirped the blue bird of the microblogging service Twitter, while a little white alien with antennae accompanied Reddit, a social networking service that provided online conversations for “digital natives,” as they were dubbed. And not only platforms but also technology companies seemed to have a thing for mascots, from Tux, the penguin of the Linux operating system, to the black Octocat that had landed on the 404 pages of Github, the web-based hosting service for software development projects. And there were many more, like the bare-bellied chimpanzee with a postman’s hat who helped create professional email for MailChimp; or the big-eyed brown owl that had become part of the logo of Hootsuite, a social media management dashboard; or the flying beaver that sat enthroned on the online travel page of a start-up company called Hipmunk. Even a nonmascot service like Facebook introduced a character, the Zuckasaurus, which looks “like a short Barney, the kid’s television show dinosaur” (Bilton 2014). Standing on its two feet while checking its laptop, the blue dragonlike dinosaur was first spotted in April 2014, when it started to address users in a pop-up window with the educational concern that it “just wants to make sure you’re sharing this post with the right people” (Bilton 2014). In short, animated animals could be found all over the World Wide Web as if it were a fairy tale. Mascots had spread from sports, where they were supposed to bring luck to a team, to the internet, and academic books started to analyze the phenomenon (Brown and Ponsonby-McCabe 2014). In the offline world, brands that were targeting their products to adults generally refrained from using mascots; companies that produced cars, alcohol, or even entertainment electronics rarely considered an animated animal as part of their brand strategy.

Parallel to the appearance of the online mascots, a similar development could be found on search pages: the rise of the Google Doodles, which introduced a new, unique style of commemoration that shared the same tendency. Until 2010, Google had only sporadically changed its prominent search website logo into those “Doodles” to mark an anniversary or event. Although the concept of the Doodle was born at the very start of the company (1998), when founders Larry Page and Sergey Brin changed the logo with a stick figure drawing to mark their visit of the Burning Man festival in the Nevada desert, the logo was not changed very often. It took two years before they requested a second change to honor Bastille Day, commemorating the beginning of the French Revolution each year on July 14. Before 2010, the logo was changed only on rare occasions. Then one could find a sketch that playfully intertwined the topic of an event with the logo: the birthday of English mathematician Ada Lovelace, Martin Luther King Jr. Day, or Halloween. After 2010, the frequency with which Doodles replaced the logo intensified. In 2010, Google published thirty-five Google Doodles, more than in any previous year. In the years 2011 and 2012, this number went up to seventy-six and eighty-three, respectively, and has gone up ever since. More and more Doodles displayed events or presented persons shaping human history and culture with imaginative cuteness. They started to appear worldwide, thereby taking national cultures into account: Britain celebrated the eight-hundredth anniversary of the Magna Carta (2015), Mexico the Day of the Dead (2013), and the United States the Mexican Hollywood actress Katy Jurado (2018).

Considering that Google is now an essential part of our public sphere—the Court of Justice of the European Union (2014) indicated this by its ruling that natural persons have the right to be forgotten and links to personal data must be erased in this public space—Google Doodles are the monuments we find in it. As we pass by those monuments when searching, we are reminded of important moments that have shaped our human fate. This form of commemoration, however, happens in a rather unique way, different from historic monuments cast in stone and erected on our public squares, which foster a certain symbolism and spread an air of pathos. Indeed, most public monuments in stone or bronze are slightly pathetic, from the Statue of Liberty enlightening the world from Liberty Island in Manhattan to the Soviet War Memorial in Berlin's Treptower Park to the Monument of the People's Heroes in Beijing's Tiananmen Square to Christ the Redeemer in Rio de Janeiro cresting Corcovado mountain. Online Doodle monuments, on the other hand, turn achievements into playful stories with imaginative cuteness and are supposed to be “fun” (Google Doodles Archive 2018). It should come as no surprise that they more often commemorate birthdays than deaths.

Before judging Google Doodles as “history light,” however, it is important to take a step back and get a full view of the transformation. Certainly all three developments—the rises of flat design, brand mascots, and Google Doodles—show a common tendency, as their style is equally defined by colorful surfaces, big typography, and playful stories or mascots, thereby resembling elements we are familiar with from children’s books or apps. Thus what is the specific form of interpellation that can be noticed here? How is technology addressing us? To state the obvious, online technology has started to address us as if we were children. The extent of this infantilization, however, only comes fully into view when comparing the described design tendency to an older project designed by Dieter Rams, who helped the company Braun to relaunch an educational toy called Lectron; and like many of his other designs, it became iconic.

Lectron was a modular electronic experimentation kit designed to introduce youth to basic electronic circuits and theory. From 1967 on, the German designer and his team, among them Jürgen Greubel, produced the packaging in a new style, including a redesign of all manuals. Being supervised by Rams, it is not very surprising that the Braun Lectron Hobby Set Radio Receiver (1969) is kept in a minimal style. Contrary to the users of Google’s search engine, Apple’s iPad, or the service online brands, however, it does not target adult users. As a game, it is tailored to a much younger age group. So how does Lectron approach its teenage user?

The cardboard box cover shows three photographs. Two smaller ones display the white radio set in Rams’s minimal design and a detail of a printed circuit board; the bigger photo pictures a black-haired teenager in a buttoned-up blue shirt, who sits in front of components and tools soldering electric parts. Lectron approaches

[Figure 2.1] The Hobby Set Radio Receiver design by Dieter Rams and Jürgen Greubel, 1967. Photograph by dasprogramm.

the technically interested and capable teenager. Contemporary flat design, on the other hand, incorporates design elements for a much younger age group. Its colorful surfaces, big typography, and animated characters are generally design elements used for targeting children aged two to seven—a time during which children are in the sensorimotor stage. Children in this stage, as the child psychologist Jean Piaget has shown, assign active roles to things in their environment (animism), while their activities are mainly categorized by symbolic play and manipulating symbols. It is a stage in which physical operations are more dominant than mere “mental” operations. Thus the conclusion is obvious: we are addressed by technology as very young children.

Fighting back the natural reaction to all miscategorizations (feeling insulted), this is an interesting outcome to be investigated further by shifting our attention back to the aspect Althusser had in mind when discussing being addressed as a form of power. So what is the effect of this infantilization of user interfaces? What force or form of power play are we facing here? For that we face a form of power play can almost be taken for granted—when technology is communicating with us in this way, it is surely not just transmitting the friendliness of cuddly Silicon Valley companies that commissioned plush toy–like interfaces to comfort us in the exhausting world we live in. To understand this manipulation further, the next section categorizes this infantilization.

How We Are Getting Manipulated

Technology has always manipulated us (Winner 1989, 19), and it does this more openly than ever, since it has started to speak. For this, one does not even need to turn to conversational interfaces, such as Apple’s Siri or Amazon’s Alexa, quarreling with us if the lights should be on or off. This also can be easily noticed by anyone who has been disciplined by a car’s navigation system. In fact, Global Positioning System (GPS) usage is a good example of a simple form of manipulation, as it has turned into quite a dominant system. To get their exact position, smartphones and millions of other devices use GPS, which was launched 1978 by the U.S. government. The system’s Master Control Station is located in the Schriever Air Force Base near Colorado Springs, overseeing thirty-two GPS satellites (U.S. Naval Observatory 2018). Currently only Russia operates an alternative system, GLONASS, with Europe and China working on further alternatives. But most cars and smartphone maps use the GPS signal, which is then correlated to a road or a calculated route. The route, however, does not always coincide with reality. A survey for Michelin (2013) among 2,200 U.S. drivers showed that 63 percent of those who use GPS say that it has led them astray at least once by pointing them in the wrong direction—and some of us obey those directions more than others.

In the United Kingdom, a driver continued to follow the navi's instructions, which told him the narrow, steep path he was driving on in Todmorden, West Yorkshire, was a road. He only noticed the mistake after he struck a fence and his BMW hung off the edge of a cliff. In South Brunswick, New Jersey, a driver ignored the end of a road because it was differently displayed on his navigation system. Following the navi's version of reality, he ignored a stop sign and hit a house. In Australia, three Japanese tourists drove their car into the Pacific Ocean. Their navi had told them there was a road to the North Stradbroke Island. After five hundred meters, they got stuck in the mud, their car being flooded by the tide. In Bergün, Switzerland, the navigation system told a man to turn onto a trail. The trail was for goats. The minivan that he had driven up that trail could only reach the road again with the help of a heavy-lift helicopter. In Italy, two Swedish tourists drove four hundred miles to the wrong Capri. Instead of relaxing on the island with its blue grotto, they ended up in an industrial city in Italy's northern region that bears the same name. In all cases, human judgment was distorted by technology, it seems. But the dialogue between human drivers and advising technology only looks at first sight like a master discourse, in which human servants blindly follow a directing technology. Technology, as both Simondon and Arendt have reminded us, is not necessarily an opposing force that aims to bring humans under control and is wrongly thought of through the template of master and servant. After all, in the preceding cases, the advice of technology could have easily been ignored. Thus one could also say that in most cases, the drivers, often tourists who were not familiar with their environment, followed "their" technology instead of asking other humans for help. In other words, we are part of this manipulation—and the same is the case when we look at patronizing, talkative self-service checkouts.

One of the countries in the West that embraced self-service checkouts early was the United Kingdom. By 2015, Tesco, the United Kingdom's largest supermarket chain, had already introduced twelve thousand of them. To help shoppers understand how to operate the new technology, the checkouts give verbal guidance on how to use them. And their most renowned comment in their early phase became "Unexpected item in bagging area. Remove this item before continuing." The reason for this comment: its pay mechanism has integrated scales. It weighs the item after it is placed in the grocery bag; this is done to ensure that the shopper pays for all the items in the basket. The problem is, however, that the system gets easily irritated, for example, when an item is too light and the second scale fails to recognize it. In these cases, the checkout announces loudly that there is an "unexpected item in bagging area" and soon after starts nervously flashing a light and an alarm sound for everyone to hear and see—the system calls for help, as it needs the reassurance of an assistant. Does it accuse you of being too thick to use it? Or suspect you of being a thief who has just stolen something? Being addressed by it in an Althusserian manner—"Hey you, there!"—we react annoyed. We recognize that other humans who see and hear this might put us into the category of social subjects who have problems using a self-service checkout, which is not very flattering.

Here we experience manipulation: when making you behave in the right manner or advising you to do the right thing, both the self-checkout and the car navigation assistant are forms of disciplinary manipulation, in contrast to those open forms of manipulation we find with infantilization, which do not directly tell you what to do. This seems to be of a different kind, with its interface *not* disciplining us but simply suggesting a situation. Cheerful design signals a simple and unproblematic context. By addressing us as very young children, the playful interfaces of flat design suggest that there is no need to understand anything. Just try it: go press this button, speak to it, create! The simple but colorful appearance signals that the users can be free from second thoughts about the complexity of the technological apparatuses as well as about the complexity of the world we live in.

We are manipulated into a situation we seemingly don't have to question—and this is why we should pause. For we have reached our first conclusion: having looked at how technology is addressing us, this chapter could establish that it is recruiting us as very young children. But can we really read the situation as technology concealing its mode of operation to lure us into its unquestioned usage? Would this not mean that we have positioned ourselves again in opposition to technology? After all, this chapter does not plan to study the concealed interests of technology companies. Instead, it aims to analyze and understand our *being with technology* by analyzing our current dialogue with it through looking into its actual "concretization" (Simondon 2017); indeed, Simondon discussed the intuitive approach of children toward technology as one way of understanding the being of technology: "One cannot study the status of the technical object in a civilization without taking into account the difference between the relation of this object to the adult and to the child," he writes (106). The technical training of the child is based on practicing with technology bringing forth a "technical subconscious" (107), which can also be understood as an intuitive skill. This experimental skill is a certain intuitive mode of technical knowledge also linked to "experts"; Simondon names the operational knowledge of farmers or of craftsmen about the material they work with. Their technical training consists of "intuition and purely operative schemas that are very difficult to formulate or transmit through any kind of symbolism" (107). Instead of scientific knowledge, the operational knowledge is created through technical realization:

> Technical realization, on the contrary, provides the scientific knowledge that serves as its principle of functioning, in the form of a dynamic intuition that can even be apprehended by the young child, and which is susceptible to becoming more and more elucidated, doubled by a

> discursive form of comprehension. . . . Through technics, encylopedism could thus find its place in the education of the child without requiring capacities for abstraction, which the young child does not fully have at its disposal. In this sense, the child's acquisition of technological knowledge can initiate an intuitive encyclopedism, grasped through the nature of the technical object. (124)

Following Simondon, and linking his understanding of intuitive encyclopedism to our problem if being recruited as very young children, one could therefore also understand the "call" of technology as an invitation to learn about a digital interface. We, however, read this dialogue according to the idea that technology is manipulating us into being its slave users, which seems to be a rather anthropomorphic reading of technology: it treats technology as if it were a human in the role of an acting subject. As pointed out earlier, technology has agency and is a force, but to understand the alienness of this force means to remind ourselves that it is not a human subject that follows a Hegelian interest to subjugate and control other humans.[1] Technology creates specific situations—in this we can find its force—but when creating those situations, it does not follow a specific interest, and this is exactly why Donna Haraway (1991, 161) in "A Cyborg Manifesto" sees the potential for "rearrangements in world-wide social relations tied to science and technology." What is created by technology can always be interpreted in different ways—if its force is understood. Even Marcuse (1998, 42), whose take on technology is generally rather critical, writes that "technics by itself can promote authoritarianism as well as liberty, scarcity as well as abundance, the extension as well as the abolition of toil." Technology is not neutral—its force is that it confronts us with a specific situation or a specific transformation; how this transformation is interpreted, however, and which concretization is going to appear is always adapted by us humans, as we are part of the technical ensemble. To say it with Donna Haraway: "We're living in a world of connections—and it matters which ones get made and unmade" (cited in Kunzru 1997).

Returning with this insight (that technology creates situations, although without interest) to our childish dialogue with technology, reading this dialogue through Simondon's approach of an intuitive encyclopedism, we can still find a negative effect of our infantilization: the creation of a situation that does not need to be further questioned. But can the recruitment of technology addressing us in an infantilizing manner be thought of differently? Can we move beyond the template of master and servant? To follow this question, the next section explores infantilization from a different perspective, by looking at an advertisement of the company that created the style of flat design: Microsoft.

In 2014, Microsoft aired its first national Super Bowl advertisement, a one-minute video produced mainly in-house. Using Microsoft products, it explores technology through the eyes of Steve Gleason, a former NFL player who is battling ASL, a severe illness that attacks nerve cells in the brain and spinal cord that control muscle movement. At the beginning of the video, we hear a computer-generated voice asking, “What is technology?” and see it being written by Steve Gleason, who sits in a wheelchair with a keyboard he operates via eye movements. We see a girl playing with a red windmill. From there, the commercial cuts to symbols that resemble written code, followed by Microsoft's colorful card screen design. Then a surgeon is flipping through large medical images displayed on a wall using hand gestures, followed by a white toy robot, which is about to look at us, as the camera movement suggests. Gleason's next question can be seen and heard: “What can it do?” after which a small boy enters the screen playing baseball standing on two artificial legs, followed by the ninety-eight-year-old painter Hal Lasko, partially blind, painting a colorful landscape with the help of a mouse. Again, Gleason's artificial voice is asking, “How far can we go?” We see pictures of a satellite in the universe, a surgeon using his hand to control an X-ray, and two groups of children cheering each other via a video-chat projection. After this introductory period, the next thirty seconds are grouped around a theme showing the examples of the “power” of technology, as

Gleason puts it: a soldier being remotely present during the birth of his child; a small child freaking out with joy when she sees her dad on the screen; several scientific and medical successes, from the launch of a rocket to a man with an artificial arm moving his hand and the emotional reaction of a women making remote contact with someone on the other side of the screen. It ends with the slogan “It has given voice to the voiceless,” showing Gleason in his high-technology wheelchair, a computer helping him communicate, his son on his lap, to whom he now connects directly by raising his eyebrows. The main slogan appears—“Empowering us all”—to be replaced after a few seconds by Microsoft’s logo.

The commercial is informed by the topic that frames it—how technology helps, “empowers,” those we love and care for to lead better lives—and certainly appeals to our emotions. The majority of the situations depicted in this video are related to health and science. Thus the situations visualized mainly pertain to health or science—generally areas not dominated by children. The video, however, uses nearly as many images of children (as individuals and in groups) as of adults. A content analysis[2] shows nine sequences with the focus on children and twelve with the focus on adults. The reason for images of curious, excited, and playful children lies partly in the task of every commercial: to create appealing images. But there is more to it. That children are playfully discovering technology is also symbolic. This becomes apparent when Gleason’s first question opening the video—“What is technology?”—is followed by a sequence showing a small girl in a dress curiously looking at the windmill she puts into motion with her small hand: humans exploring technology. The message of a girl putting a windmill into play (its movement enhanced by a sound effect) is visually answering this question. Moving a windmill means exploring technology. *The usage itself* is an act of exploration—and empowering.

Of course, one can argue that this is a message in the interest of Microsoft: the sheer usage of its commercial products is empowering—and not programming code yourself, as, for example, open source software would allow. Being able to understand or even program code yourself can certainly be more empowering. Still, this does not fully explain why the question “What is technology?” finds a fitting visual sequence in a child playing with a windmill. Instead of asking what a windmill has to do with digital media or Microsoft, the sequence makes sense. Linking this image to theories of learning and its role for the history of graphic user interfaces, the next section aims to explain why this could be the case.

Logic Is Not a Derivative of Language

The graphical user interface has become a commercial success, although this took several experiments, among them Douglas Engelbart’s NLS system, Ivan Sutherland’s Sketchpad, SGI’s Iris, the two interfaces of the Xerox Alto and Xerox Star, and the Apple Lisa and Apple Macintosh. As such, it is generally referred to as the transformation that helped personal computers to become mainstream (e.g., Chun 2011, 59). Its advantage: it is easier to use than a command line interface. Therefore the graphic interface appeals to users not familiar with coding. This section aims to inquire what it is that makes it easier and how this is linked to the girl playing with a windmill. To show this, it is first necessary to compare the older command line interface with the newer graphical user interface with respect to learning. In principle, both interfaces have the same function: they are ways to command a program. How they approach the user, however, is different. A graphical user interface’s windows, icons, menus, and pointer are intuitive elements, whereas the knowledge to operate the command line needs *to be learned beforehand.* A graphical user interface can be operated without much knowledge as it *incorporates the learning into its usage.* Learning theories in fact played an important role in its development. Discussing the work of mathematician Seymour Papert (1963, 1968), who collaborated closely with child psychologist Jean Piaget and also influenced the computer scientist Alan Kay, this section takes a look at the connection of learning theories to computer science in general and the graphical user interface in particular.

When developing new approaches to artificial intelligence, Papert had come across theories of learning by child psychologist Jean Piaget. The South African had met Piaget when he spent time in Paris as part of his second doctorate in St. John's College in Cambridge and decided to follow him to his Institute in Geneva to apply his theories to artificial intelligence, a field that found itself in its golden years from 1956 to 1974, driven by new discoveries and funding. More precisely, Papert's aim was to enhance machine learning by incorporating Piaget's ideas of the learning of children, although their interest was mutual: Piaget endorsed Papert's cybernetic approach and published many of his articles in his journal *Études d'Épistemologie Génétique.* Known today as a child psychologist, he understood himself as a scholar of epistemology exploring theories of knowledge with the aim to establish a new approach toward understanding. And it would be the graphical user interface that would pick up this approach to show that children's learning can indeed be applied to adults' learning too.

Interested in multiple ways of knowing, Piaget turned to children's learning as a unique form of interacting and theorizing. Curious about their thinking, he took their logical reasoning seriously, even when their thinking led to "wrong" answers. His nonjudgmental approach enabled him to describe four universal stages of cognitive development that are still relevant to contemporary psychology. More important in the context of this argument, however, is something different: central to his approach was the hypothesis that for human understanding and learning, the act of reasoning (the work of the mind) is as important as practical or experimental understanding (the work of the fingers and mind together). When observing children between the ages of two and seven, Piaget recognized a specific way in which children play. He saw in children's sensorimotor approach a form of learning—thinking with fingers—most important when we are very young children. From this, he concluded that logic is formed not only in the brain:

> I believe that logic is not a derivative of language. The source of logic is much more profound. It is the total

> coordination of actions, actions of joining things together, or ordering things, etc. This is what logical–mathematical experience is. (Piaget 1972, 13; see also Piaget 1969, 90)

Piaget developed what has come to be known as constructivism, an approach that viewed learning as a reconstruction rather than as a transmission of knowledge. It valued experience highly and understood playing—the manipulating of materials—as a way to create knowledge:

> To know an object, to know an event, is not simply to make a mental copy, or image, of it. To know an object is to act on it. To know is to modify, to transform the object, and to understand the process of this transformation, and as a consequence to understand the way the object is constructed. . . . In other words, it is a set of actions modifying the object, and enabling the knower to get at the structures of the transformation. (Piaget 1972, 20)

To apply and automate this approach to machine learning, Papert (1963) developed a project called "genetron," which explored the learning of algorithms by allowing them to build their own network topologies that simulated qualitative and quantitative developmental change (Shultz et al. 2008; Minsky and Papert 1969). He was later assisted by Marvin Minsky, with whom he cofounded MIT's Artificial Intelligence Lab. Despite support from MIT, the project struggled with technical limitations (Shultz et al. 2008). But Papert had also started to approach the relation of child and machine through another angle, manipulating not the machine's learning but children's learning. Applying Piaget's theory, the aim here was to allow a coordination of actions—acting with an object—to initiate learning in children: learning to operate a computer. Together with his colleagues Wally Feurzig and Cynthia Solomon, Papert developed LOGO, an educational dialect of the functional programming language Lisp, which was used to command first a virtual turtle, then a small turtle-shaped robot that could move and draw. And it was this approach that would inspire Papert's colleague

Alan Kay (1972) to develop a graphical user interface not just for children but also for "children of all ages."

When he met Papert, Alan Kay was a young, creative computer scientist who had thought about the graphical user interface ever since he was a student—the first thing his supervisor gave him to read was Ivan Sutherland's description of the Sketchpad, one of the first interactive computer graphics programs. But it was watching children in schools using Papert's LOGO that enabled a breakthrough:

> Here were children doing real programming with a specially designed language and environment. . . . This encounter finally hit me with what the destiny of personal computing really was going to be. Not a personal dynamic vehicle, as in Engelbart's metaphor opposed to the IBM "railroads," but something much more profound: a personal dynamic medium. With a vehicle one could wait until high school and give "drivers ed," but if it was a medium, *it had to extend into the world of childhood.* (Kay 1996, 523, emphasis added)

Kay understood that the logic of the world of childhood could be extended to adults by reapplying visual thinking to an adult interface. Reading (besides Piaget) the educationalists Jerome Bruner and Maria Montessori had convinced him that not the command line but visual thinking and a more iconic approach (531–32) would shape future ways of operating a computer. His insights culminated in his proposal "A Personal Computer for Children of All Ages" (Kay 1972), which described a portable educational computer to be commanded by experimental actions. It was based on a program that came to be known as Smalltalk, a program "environment in which users learn by doing" (547). Via Papert, Piaget's insight that logic can be a coordination of actions had found its way to Kay's interface; Kay saw Piaget's thesis confirmed: "Just *doing* seems to help" (547)—a seismic shift. With the graphical user interface, experimental thinking started to assist linguistic thinking. And with the rise of digital media, interfaces have become the way we approach information, an approach based on experimental as much as on linguistic logic. Relying on a logic we use in Western culture primarily when we are very young, interfaces address us as very young children. Users of graphical interfaces are asked to apply an experimental logic, which means to learn to understand the interface via a set of actions. Ever since the rise of digital media, the devices that inhabit our kitchens or gardens have stopped asking us to read through the manual before being switched on for the first time.

The infantilization of interfaces does not necessarily mean that technology is becoming smart while we are declared stupid. The manipulative dialogue of today's interfaces is not necessarily an act to deceive the user. Reaching out to a human logic mostly used in childhood, similar to the way Kay's and Papert's interfaces functioned, the playful addressing of the user can also be read as an invitation to experiment. In experimenting, in playing with the windmill, we use digital technology. Using it, however, means to understand how to act on it—acquire the skill to use its force—thereby entering into a dialogue with that technology. Entering into this dialogue is important not just for the case of the graphical user interface but also for artificial intelligence and machine learning, about which Shan Carter and Michael Niessen (2017) have argued that its new form of computing must be linked to a new and different interface to fully unfold its operational knowledge. To bring forth this operational knowledge in a more general sense, digital technology is calling upon us as children. It is not addressing us as adults, as engineers. To call into action an intuitive, visual-operational knowledge, marginalized in our postindustrial Western societies, it is recruiting us as children of all ages. The force of communication we face in digital technology is an operational knowledge; to make use of it, we are being framed as very young children.

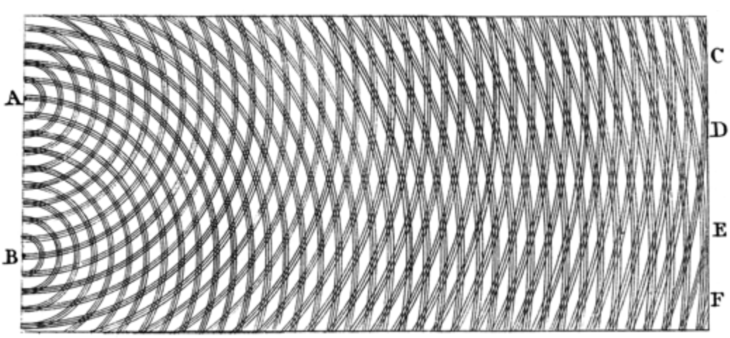
Nondialectical Dialectics

The hypothesis that digital technology finds itself linked to a specific force could be shown; still the analysis cannot stop here. For within this force, an interesting setup of power relations unfolds, power relations that are coming into action when we communicate using digital interfaces. Is the infantilization of interfaces inviting us to experiment with those interfaces, or is it luring us into a playful situation that is not to be intellectually questioned? To understand our contemporary being with technology, another effort needs to be made to explore the lines of power that run through it. How do we know if a digital interface is addressing us with the aim of empowerment, or deceiving and sedating us? How can one conceive the difference? This is the difficulty when it comes to being addressed as children: the infantilization of interfaces is able to be both patronizing *and* empowering simultaneously—the power we find within the force of communication refrains from following a well-behaved dialectical thinking.

Being patronizing *and* empowering means that one cannot be *for* or *against* infantilization. Being *for* the user's emancipation does not equal being *against* infantilization. The conceptual architecture we find at work here does not unfold in an oppositional way. An interface can be both patronizing and empowering in the same moment and is therefore not fitting into the antagonistic concept of dialectics, thesis and antithesis. Questioning the phenomenon of the infantilization of interfaces further with regard to the powers at play here, however, one also can realize that at the same time, an antagonistic, dialectic relation is not completely gone: an interface can be patronizing and empowering at the same time, although to be patronizing and to be empowering remain fundamentally different acts of power. While empowering users means that we are learning to use the power of technologies ourselves, patronizing guides and shoves us toward just acting out that power. One time the power is with the user; the other time the power is just lent to the user—in other words, there is still a fundamentally dialectic relation between. Deep inside the conceptual architecture, a negative relation, this complex force of negativity that has been described by Susan Coole (2002) and Benjamin Noys (2010) for thinking/acting difference is still at play, ensuring that there is difference.

From this follows that, again, we need to try coming to grips with the force of communication and the forms of power we find in its act of infantilizing the user. For this, the last section of this text turns to the inspiration of a visual, operational knowledge (inspired by Alan Kay and Gilbert Simondon) which it finds in the concept of “diffraction” as it appears in and has been visualized for quantum mechanics. Diffraction describes the phenomenon of waves interfering with each other, although differences remain, much like in Thomas Young’s image from 1803 (Figure 2.2) showing a two-slit diffraction.

The double-slit experiment with two waves interfering has become the thought experiment that is expressing puzzles of quantum mechanics, such as the wave–particle duality. In this century, diffraction also resurfaced as an interesting concept to think difference and was explored in depth in the writings of Karen Barad.[3] Inspired by particle diffraction of quantum trajectories, such as



[Figure 2.2.] Thomas Young’s sketch of two-slit diffraction presented to the Royal Society in 1803.

diffracted light waves, the philosopher with a doctorate in quantum physics developed the method of reading of insights through one another that came to be known as the *method of diffraction.* Barad (2007, 137) is interested in the phenomenon of diffraction as it allows her to think differences not as essentials but as a process. Diffractive patterns are always fundamentally linked to the agential apparatus that produces them, and vice versa: “Changing patterns of difference are neither pure cause nor pure effect; indeed, they are that which effects, or rather enacts, a causal structure, differentiating cause and effect.” Here I’d like to take up Barad’s aim of deessentalizing difference but to mirror and link it to the difficulties in differentiating the two modes in infantilization, that is, to be empowering and patronizing at the same time. The circumstance of infantilization’s two effects—empowering and patronizing—resembles diffraction: two waves that overlap to build a diffractive pattern. The particles/waves overlap while the waves still can be differentiated. Thus, as the image shows, despite them overlapping, there can still be difference. Or in other words, a diffractive pattern, as we find it within the phenomenon of infantilization, does not mean its effects cannot be differentiated. Following Barad further, we therefore ask the question again: how can one conceive this overlapping difference?

As Barad stresses, to understand diffraction, to know what kind of diffraction is the case, it is important to look *further* than just noticing that there is a pattern: “Crucially, diffraction effects are attentive to fine detail” (91). It is here where we find an aspect central to her approach: the detail. In her own words: “Attention to fine details is a crucial element of this methodology” (92). One has to be “sufficiently attentive to the details” and is “thinking through the details” (73), because “fine-grained details matter” (90). It is the “level of detail” (42) that enables one to answer a question. Thus it is to the detail she looks to situate difference: “Small details can make profound differences” (92). While the interference of the waves is a given—otherwise, there would be no diffraction—the way a diffraction pattern looks can vary as it is linked to its parameters: “If any of these parameters is changed, the pattern can be significantly different" (91). Only when looking at the details of the pattern and studying the "concrete" effects does one understand what exactly has been produced and which tendency of both—empowering or patronizing—precedes.

Unsurprisingly, pointing out those ambiguities and exploring their details also has become a habit of media and technology scholars interested in describing social formations. For this, theorists of digital technology and media have questioned word pairs like public–private, global–local, free–controlled, nature–technology, and work–play. Once understood as antithetical, they have made clear that their conceptual relation does not seem to be essentially oppositional anymore. Tiziana Terranova (2004) was among the first to discuss the ambiguity of work–play, pointing out that commenting online on platforms is free labor playing in the hands of companies looking for profit, although it remains pleasurable—a paradox. Wendy Chun (2011) also showed early that digital media is spreading democratic freedom along with the fact that it also accelerates the potential for global surveillance—an observation she later extended into digital media entering our daily habits, thereby messing "with the distinction between publicity and privacy, gossip and political speech, surveillance and entertainment, intimacy and work, hype and reality" (Chun 2016, ix). Analyzing algorithmic security practices and data technologies, Claudia Aradau and Tobias Blanke (2018) have disclosed how the dichotomies of normality–abnormality, friend–enemy, and identity–difference have been fundamentally reconfigured. Looking at the matter of media, Jussi Parikka (2015) dissects the opposition of nature–technology, which brings out the dependency of today's media from nature (Parikka 2015). Traversing computer science with a philosophical perspective, Luciana Parisi (2015) has questioned today's critique of instrumental rationality, pointing out that incomputability and randomness need to be conceived as the very condition of computation and not instrumentality. Pointing out dependence in a networked age, Anna Watkins Fisher (2016) discusses interventions of corporations like Walmart or McDonalds, which aimed to help their employees master problems created through being exploited by the very same corporations. One could add Nicole Starosielski (2015), Christopher Kelty (2012), N. Katherine Hayles's (2017) study of the cognitive nonconscious, and many more whose recent books or essays discuss how to deal with the ambiguities of new media and the paradoxes we live with—the force digital technology confronts us with.

These examples show that digital technology in the twenty-first century is characterized by a dialectical setting in which disparate aspects no longer operate in an oppositional mode, although their dialectical relation has not collapsed—one is the flip side of the other. Such a setting, in relation to the work of Pheng Cheah (2010), could be described as "nondialectical dialectics." *Nondialectical* as an interface that is addressing us as a very young child is both patronizing and empowering and *dialectic,* as both moments are still marked by an antagonistic relation, with one enabling the use of power while the other is just lending it. Thus, regarding digital technology, the task we face is to understand how to adjust the frame in a way that fortifies the waves of empowering by turning to the fine details. It is not to choose the right side.

IIIIIIIIIIIIIIIII

This chapter set out to study a force and found it linked to a figure of power that it described as "nondialectical dialectics." Interested in understanding how technology is addressing us, it aimed to explore how a specific force unfolds in digital communication. Drawing on Althusser's theory of interpellation, it identified a particular situation opening up when being addressed by digital technology communicating with us: digital interfaces, which aim to reach a general user, show a tendency of infantilization. By drawing on design elements from a child's world, such as big typography, primary colors, big buttons, and animated mascots, those interfaces are addressing their users as young children, thereby calling upon an experimental–operational knowledge rather than an encyclopedic–scientific one. This type of knowledge, as could be shown, has also historically been at the core of the development of graphical user interfaces, which Alan Kay or Samuel Papert conceptualized and built, inspired by the educational research of Jean Piaget, who believed that the coordination of actions ordering and joining things together should also be understood as "logical–mathematical experience."

In this operational dialogue with digital technology, however, a new phenomenon could be seen: it is not in a strict sense defined by a dialectical logic of right or wrong dialogues with technology—and in this lies the political sticking point. An interface that invites us to an experimental dialogue exploring it can be empowering, while it is not far from an interface that simply suggests how to use it best without the user gaining any deeper knowledge about it (but getting things done quickly). In other words, advising interfaces that address us as children *can* but *do not have to* be empowering—the force of digital technology that came into view could and does go both ways. The cases analyzed here, from historic Google Doodles to flat, colorful buttons on touch screens, are examples of infantilization that show that the way digital technology is addressing us is deeply ambiguous. Digital technology can produce two or more antagonistic effects at the same time and can therefore be described as being nondialectical. Still, a dialectic relation remains, as the effects it produces can be considered antagonistic with one being the flip side of the other. Only when turning to the details (Barad 2007), only when analyzing the actual effects, can the actual political scale be understood.

The force of communication that then comes into view is a complicated, ambiguous one. It is a challenge—a challenge because it is nondialectical while producing political effects; a challenge because it has agency but is not an acting subject. When thinking the force of digital technology, it helps to avoid understanding it in an anthropomorphic way and to instead call upon its alien logic. So I end this text with seconding what Finn Brunton pointed out in the first chapter, who was preparing us for an alien dialogue in which we find ourselves always already.

Notes

Without Wendy Chun's invitation and feedback on this contribution to, first, the Terms of Media II conference at Brown University and then to this volume, this text would not exist. Indeed, the text owes a lot to her encouragement here (and in other situations). I also owe warm thanks to the inspiration I got from the work and conversations with Finn Brunton and his aliens, waving to us through his text if one squints a little. Special thanks then go out to Paula Bialski, Goetz Bachmann, and Boris Traue for their thoughtful, informed, and thorough editorial reading of the manuscript, which improved it significantly. And thanks to the gifted Robert Ochshorn for sharing my serious interest in interfaces. Finally, I thank Michael Dieter and David Berry, whose invitation to contribute to their 2015 reader *Postdigital Aesthetics: Art, Computation and Design* (2015) gave me a first chance to grasp the idea of infantilization of digital interfaces. I am still surprised to find them sharing my perspective, the first time I presented it, which was the start that allowed me to build on it.

1 Understanding technology as a subject seems to be a projection linked to Finn Brunton's observation that human communication with aliens in space is imagined along the lines of a nonhuman agency with which we are familiar.

2 The analysis did not count individuals. Every time a new or a different sequence was introduced, it looked if the focus was on "adult" or "child," whereby groups counted the same as individuals. Three scenes were mixed. When the child plays football surrounded by a group of adults, the focus is mainly on the child (counted as child). The child birth in the surgery theater shows first adults at work; from there the camera moves to the child who was just born (counted as adult and child). The last scene shows Steve Gleason looking at the son on his lap (counted as adult and child).

3 Interestingly, Barad's strong focus on "interference" observed in the phenomenon of diffraction is somewhat close to Gilbert Simondon's approach, whose focus on the "ensemble" of technology and human—their interference—was discussed by describing the "technical reality" as one (Simondon 2017, 53). It has often been said (e.g., Combes 2013, 57) that Simondon's description of technology as an interference is informed by his concept of "individuation," which describes the process that produces an individual, although this individual is only a temporary instability—a theory he develops among others inspired by quantum and wave mechanics (Simondon 1992, 304), much like Barad. Therefore it comes as no surprise that Barad, with a doctorate in quantum physics, starts her point of departure—the preface of her book—from a very similar point of view. She writes, "Individuals do not preexist their interactions; rather, individuals emerge through and as part of their entangled intra-relating." Furthermore, she points out, "existence is not an individual affair" (Barad 2007, ix).

References

Althusser, Louis. 2014. *On the Reproduction of Capitalism: Ideology and Ideological State Apparatuses.* London: Verso Books.

Aradau, Claudia, and Tobias Blanke. 2018. "Governing Others: Anomaly and the Algorithmic Subject of Security." *European Journal of International Security* 3, no. 1: 1–21.

Arendt, Hannah. (1958) 1998. *The Human Condition.* Chicago: University of Chicago Press.

Barad, Karen. 2007. *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning.* Durham, N.C.: Duke University Press.

Bilton, Nick. 2014. "Facebook's New Privacy Mascot: The Zuckasaurus." *New York Times,* May 23. Accessed April 10, 2018. http://bits.blogs.nytimes.com/2014/05/22/a-blue-dinosaur-becomes-a-facebook-ambassador-for-1-28-billion-people/.

Bratton, Benjamin H. 2016. *The Stack: On Software and Sovereignty.* Cambridge, Mass.: MIT Press.

Brown, Stephen, and Sharon Ponsonby-McCabe. 2014. *Brand Mascots and Other Marketing Animals.* New York: Routledge.

Bunz, Mercedes. 2015. "School Will Never End: On Infantilization in Digital Environments Amplifying Empowerment or Propagating Stupidity?" In *Postdigital Aesthetics: Art, Computation, and Design,* edited by D. Berry and M. Dieter, 191–203. Basingstoke, U.K.: Palgrave Macmillan.

Bunz, Mercedes, and Graham Meikle. 2018. "Speaking Things." In *The Internet of Things,* 45–67. Cambridge: Polity Press.

Carter, Shan, and Michael Niessen. 2017. "Using Artificial Intelligence to Augment Human Intelligence." *Distill,* December 4. Accessed April 10, 2018. https://distill.pub/2017/aia/.

Cheah, Pheng. 2010. "Non-dialectical Materialism." In *New Materialism: Ontology, Agency, and Politics,* edited by Diana Coole and Samantha Frost, 70–91. Durham, N.C.: Duke University Press.

Chun, Wendy Hui Kyong. 2004. "On Software or the Persistence of Visual Knowledge." *Grey Room* 18: 26–51.

Chun, Wendy Hui Kyong. 2008. "On 'Sourcery,' or Code as Fetish." *Configurations* 16, no. 3: 299–324.

Chun, Wendy Hui Kyong. 2011. *Programmed Visions: Software and Memory.* Cambridge, Mass.: MIT Press.

Chun, Wendy Hui Kyong. 2016. *Updating to Remain the Same: Habitual New Media.* Cambridge, Mass.: MIT Press.

Coole, Diana. 2002. *Negativity and Politics: Dionysus and Dialectics from Kant to Poststructuralism.* New York: Routledge.

Combes, Muriel. 2013. *Gilbert Simondon and the Philosophy of the Individual.* Cambridge, Mass.: MIT Press.

Court of Justice of the European Union. 2014. Judgment in Case C-131/12 Google Spain SL, Google Inc. v Agenda Espanola de Protecci6n de Datos, Mario Costeja Gonzalez. Press Release No. 70/14. Luxembourg, May 13.

Crawford, Kate, and Vladan Joler. 2018. “Anatomy of an AI System: The Amazon Echo as an Anatomical Map of Human Labor, Data and Planetary Resources.” AI Now Institute and Share Lab, September 17. https://anatomyof.ai/.

Derrida, Jacques. (1977) 1988. “Signature Event Context.” Translated by Samuel Weber and Jeffrey Mehlman. In *Limited Inc.,* 1–23. Evanston, Ill.: Northwestern University Press.

Economist. 2011. “Difference Engine: The iPad’s Third Coming.” December 2. https://www.economist.com/babbage/2011/12/02/difference-engine-the-ipads-third-coming.

Galloway, Alexander R. 2012. *The Interface Effect.* Cambridge: Polity Press.

Gitelman, Lisa. 2013. *Raw Data Is an Oxymoron.* Cambridge, Mass.: MIT Press.

Google Doodles Archive. 2018. “About.” https://www.google.com/doodles/about.

Googleblog. 2010. “The Google Design, Turned Up a Notch.” May 5. Accessed April 10, 2018. https://googleblog.blogspot.co.uk/2010/05/google-design-turned-up-notch.html.

Haraway, Donna. 1991. “A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century.” In *Simians, Cyborgs, and Women: The Reinvention of Nature,* 149–81. New York: Routledge.

Haraway, Donna. 1997. *Modest_Witness@ Second_Millennium.FemaleMan_Meets_Onco-Mouse: Feminism and Technoscience.* New York: Routledge.

Hayles, N. Katherine. 2017. *Unthought: The Power of the Cognitive Nonconscious.* Chicago: University of Chicago Press.

Heidegger, Martin. (1954) 1977. “The Question Concerning Technology.” Translated by William Lovitt. In *“The Question Concerning Technology” and Other Essays,* 3–35. New York: Harper and Row.

Holt, Douglas. 2004. *How Brands Become Icons: The Principles of Cultural Branding.* Boston: Harvard Business School Press.

Iliades, Andrew. 2015. “Two Examples of Concretization.” *Platform: Journal of Media and Communication Volume* 6: 86–95.

Intacto. 2013. Homepage. Accessed April 10, 2018. http://www.flatvsrealism.com/.

Kay, Alan. 1972. “A Personal Computer for Children of All Ages.” Paper presented at the ACM national conference, Boston, August.

Kay, Alan. 1996. “The Early History of Smalltalk.” In *History of Programming Languages—II,* edited by Thomas J. Bergin and Richard G. Gibson, 511–98. New York: ACM Press.

Kelty, Christopher. 2012. “From Participation to Power.” In *The Participatory Cultures Handbook,* edited by Aaron Delwiche and Jennifer Jacobs Henderson, 22–31. New York: Routledge.

Kittler, Friedrich. 1990. *Discourse Networks 1800–1900.* Stanford, Calif.: Stanford University Press.

Kunzru, Hari. 1997. “You Are Cyborg.” *Wired* 5, no. 2: 1–7. Accessed April 10, 2018. http://www.wired.com/1997/02/ffharaway/.

Langdon, Winner. 1986. *The Whale and the Reactor: A Search for Limits in an Age of High Technology.* Chicago: University of Chicago Press.

Levine, Rick, Christopher Locke, David Searles, and David Weinberger. 2000. *The Cluetrain Manifesto: The End of Business as Usual.* New York: Basic Books.

Mackenzie, Adrian. 2017. *Machine Learners: Archaeology of a Data Practice.* Cambridge, Mass.: MIT Press.

Marcuse, Herbert. 1998. “Some Social Implications of Modern Technology.” In *Technology, War, Fascism,* 39–65. New York: Routledge.

Minsky, Marvin, and Seymour Papert. 1969. *Perceptrons: An Introduction to Computational Geometry.* Cambridge, Mass.: MIT Press.

Millman, Debbie. 2012. *Brand Bible: The Complete Guide to Building, Designing, and Sustaining Brands.* Beverly, Mass.: Rockport.

Noys, Benjamin. 2010. *Persistence of the Negative: A Critique of Contemporary Continental Theory.* Edinburgh: Edinburgh University Press.

Parikka, Jussi. 2015. *A Geology of Media.* Minneapolis: University of Minnesota Press.

Parisi, Luciana. 2015. “Instrumental Reason, Algorithmic Capitalism, and the Incomputable.” In *Alleys of Your Mind: Augmented Realities and Its Traumas,* edited by Matteo Pasquinelli, 125–37. Lüneburg, Germany: meson press.

Papert, Seymour. 1963. “Études comparee de l’intelligence chez l’enfant et che le robot.” *Études d’Épistemologie Génétique* 15: 131–94.

Papert, Seymour, with Guy Cellérier and Gilbert Voyat. 1968. *Cybernétique et Épistémologie.* Études d’Épistemologie Génétique 22. Paris: Presses universitaires de France.

Piaget, Jean, and Bäbel Inhelder. 1969. *The Psychology of the Child.* London: Routledge and Kegan Paul.

Piaget, Jean. 1972. “Development and Learning.” In *Piaget Rediscovered,* edited by Richard Ripple and Verne Rockcastle, 7–20. Ithaca, N.Y.: Cornell University Press.

Raley, Rita. 2016. “Algorithmic Translations.” *New Centennial Review* 16, no. 1: 115–37.

Rawsthorn, Alice. 2010. “Google’s Doodles.” *New York Times,* March 24. Accessed April 10, 2018. http://www.nytimes.com/2010/03/14/t-magazine/02talk-rawsthorn.html.

Shannon, Claude, and Warren Weaver. 1949. *The Mathematical Theory of Communication.* Chicago: University of Chicago Press.

Shultz, Thomas R., William C. Schmidt, David Buckingham, and Denis Mareschal. 2008. “Modeling Cognitive Development with a Generative Connectionist Algorithm.” In *Developing Cognitive Competence: New Approaches to Process Modeling,* edited by Tony J. Simon and Graeme S. Halford, 205–61. Hove: Lawrence Erlbaum.

Simondon, Gilbert. 2017. *On the Mode of Existence of Technical Objects.* Minneapolis, Minn.: Univocal Press.

Simondon, Gilbert. 1992. “The Genesis of the Individual.” In *Incorporations,* edited by Jonathan Crary and Sanford Kwinter, 297–319. New York: Zone Books.

Simondon, Gilbert. 2012. “Technical Mentality.” In *Gilbert Simondon: Being and Technology,* edited by Arne Boever, Alex Murray, Jon Roffe, and Ashley Woodward, 1–15. Edinburgh: Edinburgh University Press.

Starosielski, Nicole. 2015. *The Undersea Network.* Durham, N.C.: Duke University Press.

Terranova, Tiziana. 2004. *Network Culture: Politics for the Information Age.* London: Pluto Press.

Tesco PLC. 2015. “New Audio Voice for Self Service Checkouts.” Accessed April 10, 2018. https://www.tescoplc.com/news/news-releases/2015/tesco-to-end-unexpected-item-in-the-bagging-area.

U.S. Naval Observatory. 2018. “USNO GPS Timing Observations.” Accessed April 10, 2018. http://tycho.usno.navy.mil/gps.html.

Watkins Fisher, Anna. 2016. “User Be Used: Leveraging the Play in the System.” In *New Media, Old Media,* 2nd ed., edited by Wendy Hui Kyong Chun, Thomas Keenan, and Anna Watkins Fisher, 287–300. New York: Routledge.

Williams, Raymond. 1985. *Keywords: A Vocabulary of Culture and Society.* Oxford: Oxford University Press.

Wingfield, Nick. 2012. “Microsoft Drops Metro Name for New Product Look.” *New York Times,* August 3. Accessed April 10, 2018. http://bits.blogs.nytimes.com/2012/08/03/microsoft-drops-metro-name-for-new-product-look.

Winner, Langdon. 1989. *The Whale and the Reactor: A Search for Limits in an Age of High Technology.* Chicago: University of Chicago Press.

Symbols, Patterns, and Behavior: Towards a New Understanding of Intelligence

Paper by
Rolf Pfeifer

Formerly, professor of computer science at the Department of Informatics University of Zurich, and director of the Artificial Intelligence Laboratory.

Introduction by
Agnes Cameron

Hardware and software developer, with an interest in complex systems and simulation. Agnes is currently in residence at Somerset House Studios.

Pfeifer, R.. "Symbols, patterns, and behavior: towards a new understanding of intelligence." (1999).

Introduction

Rolf Pfeifer's Symbols, Patterns and Behaviour: Towards a New Understanding of Intelligence is a paper that has had a strong influence on how I think about 'artificial intelligence'. In it, Pfeifer makes a case that what we observe as intelligent behaviour (or not) is a function of perspective, and that all behaviour is fundamentally defined at the interface between an agent and its environment.

The title invokes 3 contrasting models by which we understand intelligence: that it is based on symbol-processing capacity (e.g. early chess computers), that it relies on pattern recognition (e.g. neural networks), or that it is, as Pfeifer contends, ultimately located somewhere in the relationship of a body with the wider world.

The paper gives a class of principles for designing agents that exhibit this embodied intelligence, placing an emphasis on autonomy, situatedness (agents control their interactions with the environment), and tight sensory-motor co-ordination (rather than viewing sensing and acting as separate 'modules' controlled by a central processor). Their final principle, that of 'good design is cheap' is one that feels particularly apt for a discussion around intelligence and interfaces:

"Leg coordination in insects does not require a central controller. There is no internal process corresponding to global communication between the legs... but there is global communication between all the legs, namely through the environment. It is mediated by a physical process, not by an information process (or a process of signal transfer) within the agent. If the insect lifts one leg, the force on all other legs is changed instantaneously because of the weight of the insect."

This example - of an intelligence wholly dependent on an interface with the environment, where apparently complex behaviour is undergirded by simple physical processes not requiring "unnecessary neural substrate" (in later work, this is termed "morphological computation" (1)) - runs counter to models that emphasise information-processing as the key to intelligent systems.

Despite being written in 1996, many of the arguments in the paper ring true today. For example, the problem of 'symbol grounding' - the idea that in order for an agent to deal effectively with abstract concepts, they must be grounded in that agents' interaction with the real world - is one we still encounter with pattern-matching machine learning systems. Given recent accidents involving self-driving cars, where the vehicles fail to adequately apply human labels to an unexpected and shifting set of circumstances (2), and still struggle in sensorially confusing conditions like rain and snow (3), it's interesting to think about how a more embodied approach to intelligence would change the way we talk about these problems.

1 Rolf Pfeifer; Fumiya Iida; Gabriel Gómez (2006). Morphological computation for adaptive behavior and cognition. , 1291(none), 0–29. doi:10.1016/j.ics.2005.12.080

2 Timothy B Lee, How terrible software design decisions led to Uber's deadly 2018 crash, https://arstechnica.com/cars/2019/11/how-terrible-software-design-decisions-led-to-ubers-deadly-2018-crash/

3 Van Brummelen, Jessica; O'Brien, Marie; Gruyer, Dominique; Najjaran, Homayoun (2018). Autonomous vehicle perception: The technology of today and tomorrow. Transportation Research Part C: Emerging Technologies, (), S0968090X18302134–. doi:10.1016/j.trc.2018.02.012

Symbols, patterns, and behavior: towards a new understanding of intelligence

Rolf Pfeifer
AI Lab, Computer Science Department
University of Zurich, Winterthurerstrasse 190
CH-8057, Zurich, Switzerland
phone: +41-1-257 43 20/31; fax: +41-1-363 00 35
pfeifer@ifi.unizh.ch

Abstract

Artificial Intelligence (AI) has two main goals, namely to understand intelligence and to develop systems which behave in intelligent ways. The classical approach views an intelligent agent as a symbol processor: it receives input from the environment, processes the information, i.e. manipulates symbols, and produces some output. This way of looking at intelligence has also been called the information processing approach. Recently symbol processing models have been criticized in a number of ways. It has been argued that they lack robustness, that they cannot perform in real time, that learning is mostly ad hoc and not performed in a principled way, and that due to their discrete and sequential nature they are more like digital computers rather than like brains. Moreover, it has been proposed that classical AI models suffer from a number of fundamental problems, ("symbol grounding", "frame problem", lack of "situatedness"). In this paper these problems will be reviewed and illustrated. It will be discussed to what extent connectionist models solve the problems of classical AI. It will be argued that, although they do solve some of them, they are not sufficient to answer the fundamental problems. In order to achieve the latter, it is necessary to study embodied autonomous agents which interact on their own with their environments. If we want to build interesting agents we have to observe a number of design principles. These principles will be outlined. They will be used to contrast the new view of intelligence with the traditional one.

1 Introduction

Artificial Intelligence (AI) has two main goals, namely to understand intelligence (or intelligent behavior) and to develop systems which behave in intelligent ways. In this paper the focus is on understanding intelligence, i.e. we adopt a cognitive science perspective. However, we deal with both issues. What makes the methodology of AI so productive is the synthetic character: if we are to design intelligent systems we have to understand behavior, and actually building systems helps us understand behavior in new ways. The cognitive science perspective implies that our main goal is to learn about intelligence. This goal also provides an evaluation criterion. Thus, we might prefer a program for playing chess over another one, even though its performance is worse. The program that plays worse might instantiate some psychological principles of perception and human memory that we consider important, whereas the winning program might simply be based on search.

Before we go on, we have to say what we mean by intelligence. It would be hopeless to try and define it—we are not very likely to achieve agreement. It is highly subjective and strongly depends on our expectations. If I, as an adult, play chess, nobody is very impressed. However, if a two year old played exactly like me, we would be very impressed, even though I am only a very average player. Rather than pursuing the question what intelligence *is*, we propose to replace the question by a different, more productive one: Given some behavior that we find interesting, how does it come about? What are the underlying mechanisms? If we pursue this line, we no longer have to argue whether ants are intelligent or not. We either find their behavior interesting and then it is worthwhile trying to work out the mechanisms, or we don't—and then we might not be interested in the mechanisms either.

Roughly the field of AI can be divided into three main approaches or paradigms, symbol processing AI, connectionism, and "New AI". This categorization also corresponds to a historical development with paradigm shifts in between. Symbol processing AI is based on the idea that intelligence can be viewed as the manipulation of abstract symbols. Connectionism, also called "neural networks", or "parallel distributed processing", refers to a particular type of modeling approach that is vaguely inspired by brain-like systems. "New AI"—or behavior-based AI—studies systems which are physically embodied and which have to interact on their own with the real world.

In the early days of AI—during the symbol processing period—the main interest was in thinking, reasoning, problem solving, language, i.e. in "high-level" human capabilities. Over time, the interest has shifted towards more simple or "low-level" kinds of behavior that relate more to sensory-motor capacities like perception and object manipulation. Connectionism has been strongly focusing on this area. More recently there has been yet another shift of interest, namely from systems that excel at one particular task to systems capable of performing many different tasks like navigating in an unpredictable world, learning categorizations in a real-world environment, collecting and manipulating objects, while maintaining battery level and physical integrity, etc. From these developments a new understanding of intelligence has emerged.

Ultimately, we are interested in developing a "theory of intelligence". Given the state-of-the-art it is entirely open what this theory will look like. In classical AI it was suggested that the theory be based on the idea of symbol processing (e.g. Newell, 1990). Connectionists believe that it will be based on parallel distributed processing of patterns of activation (e.g. Rumelhart and McClelland, 1986). Recently, it has been suggested that the mathematical theory of non-linear dynamics might provide the right framework because it is capable of capturing not only aspects of the control architecture, but of the complete system, including its physics (e.g. Beer, in press; Steinhage and Schoener, in press). Another group of researchers is capitalizing on evolutionary considerations (see, e.g. Harvey et al., in press, for a review). Principles of micro-economics have also been suggested as a framework to understand intelligent behavior (e.g. McFarland and Boesser, 1993). The field is still changing rapidly and no clear winner can be foreseen. Therefore, we have chosen to capture some of the insights gained in recent years in the form of a set of compact *design principles*, rather than in the framework of a rigid formal theory.

We will proceed as follows. First we will outline the classical approach. This will be very brief, assuming that everyone is familiar with it. Then we will point out some of the problems that eventually lead to a paradigm shift. We then discuss in what ways connectionism contributes to the resolution of these problems, using a number of examples. We then present some of the design principles for intelligent systems and discuss to what extent this new view resolves some of the basic issues.

But before starting we need to make a short digression. We have found that in the literature on AI and cognitive science there is a lot of confusion about the so-called "frame-of-reference" problem. We will give a short outline using the famous example of Simon's ant on the beach.

2 The "frame-of-reference": Simon's ant on the beach

Figure 1 shows an ant walking on the beach. The example has been taken from Herbert Simon's seminal book entitled "The Sciences of the Artificial" (Simon, 1969). Let us assume that the ant is coming from the upper right corner of the picture and walking towards the lower left one, where its nest is located. The path of the ant has been marked by a line, its trajectory. The trajectory is highly complicated because the beach is full of pebbles, rocks, and other obstacles. However, this complexity is in the eye of the observer, rather than in the ant itself. Surely, the trajectory is not stored in the ant's head and so its behavior cannot be based on it. In other words, there is no trajectory functioning, say, like a plan. Rather, the mechanisms that are driving the ant's behavior may be very simple. They might be described as rules like "if obstacle sensor on left is activated, turn right" (and vice versa). These "rules" are implemented in terms of neural structures within the ant. The neural structures are embedded in the body of the ant. The interaction of the neural substrate with the environment is mediated through the body. In this interaction with the environment the apparent complexity of the trajectory emerges. We say apparent complexity, because the complexity is in the eye of the observer, rather than being a property of the agent itself.

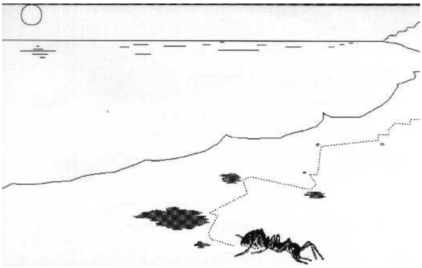


Figure 1: Simon's ant on the beach.

The trajectory we observe, refers to behavior. Behavior is always an *interaction of an agent with its environment*. It is neither a property of the agent itself, nor a property of the environment alone. The behavior is to be clearly distinguished from the internal (neural) mechanism that is responsible for it. In other words, behavior cannot be reduced to internal mechanism. Doing so would constitute a *category error*.

This seems almost trivially obvious. But if it is so evident, it is even more surprising that there is an enormous confusion about this problem in the entire literature. Throughout this paper we will refer to the "frame-of-reference" problem (for a detailed discussion, see Clancey, 1991).

3 Symbols: The traditional AI approach

3.1 Characterizing symbol processing

Assuming that everyone is familiar with symbol processing AI (sometimes called "classical AI" or "traditional AI"), our introduction will be very short. Symbol processing AI is based on the idea that intelligence can be viewed as the manipulation of abstract symbols. Newell and Simon (1976) proposed the so-called "Physical Symbol Systems Hypothesis" which, in essence, states that a necessary and sufficient condition for general intelligent action is that it be a physical symbol system. The term "physical" refers to the idea that symbol systems must be realized in some physical medium (paper, computer, brain) but it is irrelevant *how* they are realized. Typical examples of artificial physical symbol systems are production systems (or rule-based systems) or general purpose programming languages like Lisp or C. Necessary means that any system lacking this property cannot be intelligent, sufficient implies that a system having this property has the potential for intelligent action. A (symbolic) representation (figure 2) in the sense of Newell refers to a situation in the outside world and obeys the "law of representation", namely:

$$\text{decode}[\text{encode}(T)(\text{encode}(X))] = T(X),$$

where X is the original external situation and T is the external transformation (Newell, 1990, p. 59). There is an encoding as well as a decoding function for establishing a mapping between the outside world and the internal representation.

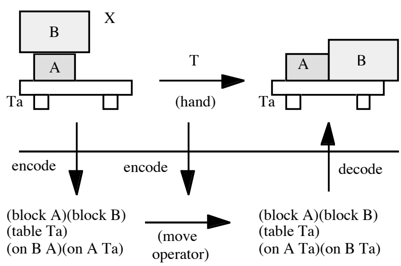


Figure 2: The "law of representation" (following Newell, 1990).

The symbol processing approach views an intelligent agent as an information processor which receives input from the environment, processes the information, i.e. manipulates symbols, and produces some output. Therefore it has also been called the *information processing approach*. Over the years, it became clear that this approach is suffering from a number of problems, which in turn lead to a paradigm shift. Let us look at some of them.

3.2 Problems of traditional symbol processing models

It has been argued that symbol processing models lack robustness (i.e. fault and noise tolerance, as well as the capacity to generalize), that they cannot perform in real time, that learning is mostly ad hoc and not performed in a principled way, and that due to their discrete and sequential nature they are more like digital computers rather than like brains. Moreover, it has been argued that classical AI models suffer from the fundamental problem of "symbol grounding" and the "frame problem", and that they lack the property of "situatedness". In the paper these problems will be reviewed and illustrated.

The well-known problems

Robustness: One symptom traditional AI symbol processing systems suffer from is lack of robustness which means that they lack noise and fault-tolerance, and that they cannot behave appropriately in new situations. Standard symbol processing models are neither noise nor fault tolerant *unless* there is explicit provision for noise and particular types of faults. The most important point concerning robustness, however is the inability to perform appropriately in novel situations, i.e. the lack of generalization capacity. If a situation arises which has not been predefined a traditional symbol processing model will break down[1].

Performance in real time: It has turned out that when systems which are based on symbol processing models are embedded in real robots they are typically too slow, i.e. they are not capable to meet real time processing demands. The reasons for this will be discussed in detail below.

Integrated learning: Traditional AI models have often been criticized because their learning mechanisms are ad hoc and imposed on top of non-learning systems. In contrast, the brain is a system which continuously learns. Humans, for example, learn always whether they like it or not. If you (the reader) read this paper you will learn something whether you like it or not, or whether you find it is useful or not. There are exceptions of classical systems where learning is an integral part of the architecture and takes place continuously like SOAR (Laird et al., 1987) but they are not representative of the majority of approaches. Moreover, SOAR, like other classical models, suffers from the fundamental problems of symbol processing systems (see below).

Sequential nature of programs: One main point of criticism has been that the architecture of today's AI programs is sequential and they work on a step-by-step basis. By contrast the human brain is massively parallel with activity in many parts of the brain at all times. Moreover, it is hard to imagine that something like symbols would be "found" in the brain. This problem is induced by the fact that current computer technology is largely based on architectures of the von Neumann type which is, at the information processing level, a sequential machine. The notion of computation abstracts from the physical realization and only considers the algorithmic level. In cognitive science the cognitivistic position makes a similar kind of abstraction: intelligent function or cognition can be studied at the level of algorithms, the physical realization of the algorithm does not matter (Putnam, 1975). We will argue later on that the physical realization indeed does matter. But we have to extend our perspective to the agent as a whole.

The criticisms of AI models presented so far are well-known. Since the mid-eighties a number of additional ones have been raised, pertaining to fundamental issues. It has been argued that traditional symbol-processing AI models suffer from the "frame problem" and the problem of "symbol grounding", and that they lack the property of "situatedness". These problems will now be reviewed in turn.

---

[1]There are some approaches in symbol processing AI (in the field of machine learning) which in fact do generalize to some extent but the ways in which generalization is achieved is typically ad hoc. See also below: "Integrated learning".

The fundamental problems

Traditionally AI models have been conceived primarily for artificial, virtual or formal worlds. Examples are search, formal games like checkers or chess, and theorem provers[2]. Whenever dealing with the *real* world two important aspects must be taken into account: (i) models must somehow relate to the outside world (otherwise there would be no point in building them), and (ii) the real word, in contrast to a virtual one, is constantly changing, intrinsically unpredictable and only partially knowable. The import of this real-world perspective can hardly be overestimated. It is at the heart of the fundamental problems.

The symbol grounding problem: The symbol grounding problem relates to aspect (i). It refers to the problem of how symbols acquire meaning. In AI the meaning of symbols is typically defined in a purely syntactic way by how they relate to other symbols and how they are processed by some interpreter (Newell and Simon, 1976; Quillian, 1968). The relation of the symbols to the outside world is rarely discussed explicitly. This position not only pertains to AI but to computer science in general. Except in real-time applications the relation of symbols to the outside world is never discussed. The—typically implicit—assumption made is that the potential users will know what the symbols mean (e.g. the price of a product stored in a data base). Interestingly enough this idea is also predominant in linguistics: it is taken for granted that there is some kind of correspondence between the symbols or sentences and the outside world. The study of meaning then relates to the translation of sentences into some kind of logic-based representation whose semantics is clearly defined (Winograd and Flores, 1986, p. 18). This position is acceptable as long as there is a human interpreter and it can be safely expected that he is capable of establishing the appropriate relations to some outside world: the mapping is "grounded" in the human's experience of his or her interaction with the real world.

However, once we remove the human interpreter from the loop, as in the case of autonomous agents, we have to take into account that the system needs to interact with the environment on its own. Thus, if there are symbols in the system, their meaning must be *grounded* in the system's own experience in the interaction with the real world. Symbol systems in which symbols only refer to other symbols are not grounded because the connection to the outside world is missing. The symbols only have meaning to a designer or a user, not to the system itself. It is interesting to note that for a long time the symbol grounding problem has not attracted much attention in AI or cognitive science—and it has never been an issue in computer science in general. Only with the renewed interest in autonomous robots it has come to the fore. This problem has been discussed in detail by Harnad (1990). It will be argued later that the symbol grounding problem is really an artifact of symbolic systems and "disappears" if a different approach is used.

[2]The field of robotics is an exception.

Situatedness: The concept of "situatedness" (or "situated cognition", "situated action", "situated agents"), has recently attracted a lot of interest and lead to heated debates about the nature of intelligence and the place of symbol processing systems in studying intelligence. For example, there is a complete issue of the journal Cognitive Science dedicated to "situatedness" (Cognitive Science, 1993). "Situatedness" roughly means the following. First, it implies that the world is viewed entirely from the perspective of the agent (*not* from the observer's perspective). Second, a situated agent capitalizes on the system-environment interaction. It's behavior is largely based on the current situation rather than detailed plans. It only focuses on the relevant aspects of the situation. And third, a situated agent is not merely reactive, but brings its own experience to bear on the current situation. In other words, the behavior of a situated agent will change over time. Because of these properties, situated agents can act in real time.

The perspective of "situatedness" contrasts with classical AI where the approach has been—and still is—to equip the agents with models of their environment. These models form the basis for planning processes which in turn are used for deciding on a particular action. In this view the agent perceives a situation ("sensing"), recognizes objects, draws a number of inferences about the current situation and about the potential effects of various actions, forms a plan ("thinking"), decides on a particular action and finally performs the action ("act"). This is called a "sense-think-act" cycle. This may work in a virtual world with clearly defined situations and given operators. But even there, plan-based agents run quickly into combinatorial problems (e.g. Chapman, 1987). Moreover, since the environment is only partially knowable a complete model cannot be built in the first place. Even if only partial models are developed, keeping the models up to date requires a lot of computational resources. Inspection of the problem of taking action in the real world shows that it is neither necessary nor desirable to develop "complete" and very detailed plans and models (e.g. Winograd and Flores, 1986; Suchman, 1987). Typically only a small part of an agent's environment is relevant for its action. In addition, instead of performing extensive inference operations on internal models or representations the agent can interact with the current situation: the real world is, in a sense, part of the "knowledge" the agent needs in order to act, it can merely "look at it" through the sensors.

Traditional AI systems, and most computer systems for that matter, are not situated and there is no reason why they should be because there is always a human interpreter in the loop. However, if we are interested in building systems which act directly in the real world they will have to be situated. Otherwise, given the properties of the real world, the system will not be able to perform intelligently.

The "sense-think-act" view also suggests an information processing perspective: the input is given by the sensing, the sensory information is processed, and an output is delivered in the form of an action. This view, while appropriate for traditional computer applications, turns out to be highly inappropriate for situated agents.

The "frame problem": The "frame problem" was originally pointed out by McCarthy and Hayes (1969). It has more recently generated a lot of interest (e.g. Pylyshyn, 1987). It comes in several variations and there is not one single interpretation. The central point concerns how to model change (Janlert, 1987): given a model of a continuously changing environment, how can the model be kept in tune with the real world? Assuming that the model consists of a set of logical propositions (which essentially holds for any representation) any proposition can change at any point in time. However, the physical world is inherently constrained by the laws of physics: objects do not simply disappear, they do not start to fly without reason, etc. But ice cubes lying in the sun do disappear. Such constraints either have to be modeled explicitly or certain heuristics have to be applied. One heuristic is that we assume things do not change unless explicitly modeled. But if this latter strategy is adopted, how about a cup on a saucer when the saucer is moved? The cup will also change its position. The problem is, that there is potentially a very large number of possible inferences which can be drawn. Let us explain this using an example by Dennett (1987).

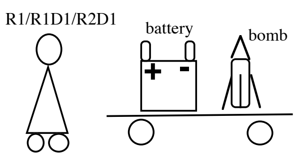


Figure 3: Illustration of the "frame problem" (following Dennett, 1987).

The robot R1 has been told that its battery is in a room with a bomb and that it must move the battery out of the room before the bomb goes off (figure 3). Both the battery and the bomb are on a wagon. R1 knows that the action of pulling the wagon out of the room will remove the battery from the room. It does so and as it is outside, the bomb goes off. Poor R1 had not realized that pulling the wagon would bring the bomb out along with the battery.

The designers realized that the robot would have to be made to recognize not just the intended implications of its acts, but also its side-effects by deducing these implications from the descriptions it uses in formulating its plans. They called their next model the robot deducer, or short R1D1 and did the same experiment. R1D1 started considering the implications of pulling the wagon out of the room. It had just finished deducing that pulling the wagon out of the room would not change the color of the room's walls when the bomb went off.

The problem was obvious. The robot must be taught the difference between relevant and irrelevant implications. R2D1, the robot-relevant-deducer, was again tested. The designers saw R2D1 sitting outside the room containing the ticking bomb. "Do something!" they yelled at it. "I am", it retorted. "I am busily ignoring some thousands of implications I have determined to be irrelevant. Just as soon as I find an irrelevant implication, I put it on the list of those I must ignore, and …" the bomb went off. (Dennett, 1987, pp. 147-148).

We have now listed the most important problems of traditional AI models. Our considerations are not restricted to AI but apply to computer systems in general. We will now discuss two solutions which have been proposed to resolve some of these issues, namely connectionism and "New AI".

4 Patterns: The contribution of connectionism

Because traditional AI was not progressing satisfactorily any more, connectionism was highly welcomed by large parts of the research community in AI. The hope was that connectionism would resolve many of the problems of traditional symbol processing AI. Indeed connectionism does contribute in interesting ways.

For example, connectionist models are fault tolerant and noise tolerant. Because of their parallel nature they preserve most of their functionality if there is noise in the data or if certain parts of the network malfunction. But more important, connectionist models have a certain ability for generalization. They are capable of behaving appropriately even in circumstances the model has not encountered before. These are the essential factors contributing to the robustness of connectionist models.

There are additional characteristics which have contributed to their popularity. Learning is intrinsic and they are in some sense more brain-like. Connectionist models consist of large numbers of nodes and connections. Typically there are too many connections to adjust manually, so they have to be tuned through learning mechanisms. Connectionist models have attracted a lot of attention since they do not only perform what has been programmed into them, but also what they have learned. In this sense they sometimes show unexpected or "emergent" behavior. In contrast to symbolic models connectionist ones integrate learning in natural ways. Examples of neural network learning will be given below.

Another aspect which has contributed to the popularity of connectionst models is more of a psychological nature. Connectionist models have been praised for being more brain-like than traditional ones (e.g. Rumelhart and McClelland, 1986). There are many units working in parallel and they process patterns rather than symbols. However, although they do draw inspiration from reflections about the brain, they reflect brain properties only in a very remote way, if at all. If we want to model real brain function, our models must look very different (e.g. Reeke and Edelman, 1989). We will not go into this aspect any further since it is highly controversial and it is somewhat marginal for the argument to be made in this article.

In summary, connectionist models are more robust, they integrate learning, and they are somewhat more "brain-like" than classical models. Because of their parallel nature, they may also cope better with real-time demands (but this latter point is debatable).

Let us now discuss to what extent connectionism contributes to resolving the fundamental issues. Connectionist models process *patterns of activation* rather than symbols. This seems a more realistic view of what is going on in the brain than the one endorsed by symbol manipulating models. Moreover, since connectionist models can learn they could potentially learn to make their own categorization of the environment, rather than having it programmed into the system by the designer. One might think that this would provide a solution to the symbol grounding problem. We will see that this is not automatically the case.

4.1 Supervised learning

Let us illustrate our argument with NETTalk, a well-known classical connectionist model (Sejnowski and Rosenberg, 1987). NETTalk translates English text into speech using a multi-layer feed-forward backpropagation network. The architecture is illustrated in figure 4. There is an input layer, a hidden layer, and an output layer. At the input layer the text is presented. There is a window of seven slots. This window is needed since the pronunciation of a letter depends strongly on the context in which it occurs. In each slot one letter is encoded. For each letter of the alphabet there is one node in each slot. Input nodes are binary on/off nodes. Therefore, an input pattern consists of seven active nodes (all others are off). The nodes in the hidden layer have continuous activation levels. The output nodes are similar to the nodes in the hidden layer. They encode the phonemes by means of a set of phoneme features. This encoding of the phonemes in terms of phoneme features can be fed into a speech generator. For each letter presented at the center of the input window—"e" in the example of figure 4—the correct phoneme encoding is known. By "correct" we mean the one which has been encoded by linguists earlier[3]. The model starts with random connection weights. It propagates each input pattern to the output layer, compares the pattern in the output layer with the correct one and adjusts the weights according to a learning algorithm, namely backpropagation. After presentation of many (tens of thousands) patterns the weights converge, i.e. the network picks up the correct pronunciation.

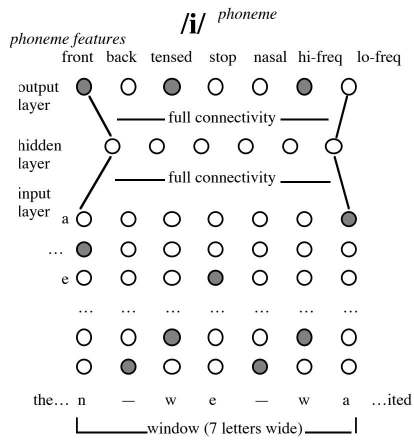


Figure 4: Architecture of the NETTalk model. Details, see text.

As demonstrated by the authors, NETTalk does indeed solve the well-known problems. "Shaking" the weights, i.e. superimposing random distortions on the weights, removing certain connections in the architecture, and errors in the encodings do not significantly influence the network's behavior. Moreover, it can handle—pronounce correctly—words it has not encountered before, i.e. it can generalize. In short, the model is *robust*. Learning is an intrinsic property of the model. Moreover, and that is one of the most exciting properties of the model, at the hidden layer certain nodes start distinguishing between vowels and consonants. In other words they are on when there is a vowel at the input, otherwise they are off. This consonant-vowel distinction has also been called "emergent".

Let us now examine the fundamental problems. Each input node corresponds to a letter. Letters are symbols, i.e. the encoding at the input layer is in terms of symbolic categories. Phoneme features are designer defined categories and thus the respective sound encodings are also symbolic. There are two points to be made. First, the system is not coupled to the environment. The interpretation of input and output is entirely up to humans who have to interpret the symbols at the input and the output. The fact that the output is fed into a speech generator is irrelevant since this has no effect on the model. Therefore, NETTalk, just like any traditional model in AI, suffers from the symbol grounding problem. It can therefore be expected that even if NETTalk is improved it will never reach a human-like performance level. It should be mentioned that the authors never claimed to be solving the symbol grounding problem with this model.

[3] In one experiment a tape recording from a child was transcribed into English text and for each letter the phoneme encoding as pronounced by the child was worked out by the linguists. In a different experiment the prescribed pronunciation was taken from a dictionary.

Second, the consonant-vowel distinction is not really emergent, but—in a sense—precoded. It can be shown (Verschure, 1992) that of those features which are used to encode vowels, only about 5% are also used to encode consonants and vice versa. In other words, the distinction is not really acquired by the system but rather (indirectly) pre-programmed by the way the examples are encoded in a symbolic way. Again, this distinction is not grounded in the model's experience but implicitly grounded in the experience of the designer of the model. Since there is no interaction with the real world (only patterns are presented to the network input layer) the frame problem and situatedness are not addressed.

From this discussion it can be concluded that supervised learning does solve a certain class of problems. But it solves the problems at an "information processing" level, rather than by interaction with the real world. Supervised learning does not resolve the issue of symbol grounding and will not lead to situated systems. The categories the model has at its disposal are given at design time once and for all. All the models can do is combine the basic categories in various ways. But the basic categories that determine how the model can interact with its environment, are fixed.

While many would probably agree that supervised models are based on designer-defined categories, there is likely to be disagreement about unsupervised models.

4.2 Unsupervised schemes

A prominent example of an unsupervised scheme is Kohonen's topological map (e.g. Kohonen, 1988a) which comes in many variations. Again, assuming familiarity, the presentation is very brief. The basic architecture is shown in figure 5. The input layer is fully connected to the map layer. In the map layer, there are lateral connections which are excitatory for close neighbors, inhibitory for those further away, and neutral for the ones yet further out. Patterns are presented to the model at the input layer and depending on the particular architecture and choice of parameters it will eventually learn a particular categorization of the input space. The details of the algorithm do not matter. What does matter is the basic principle that there is no need for the system to be given a classification of input patterns by the designer (which is why this is called unsupervised).

For example, in the "neural phonetic typewriter" (Kohonen, 1988b) the inputs are spectral patterns corresponding to pre-processed signals and the classes which are formed can be interpreted as "pseudo phonemes". Pseudo phonemes are like phonemes but they have a shorter duration (10ms rather than 40 to 400ms). In contrast to supervised learning, the designer does not predefine the classes into which the patterns need to be sorted. But does the model really acquire its own categorization in its interaction with the real world, i.e. does it really solve the symbol grounding problem? The answer is "no". Why?

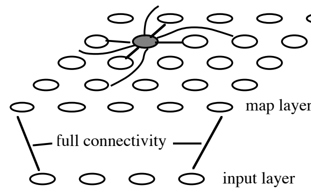


Figure 5: Basic architecture of a Kohonen network. Details, see text.

The patterns which are presented to the model have been carefully preselected by the designer. This does not imply that the designer determines the individual patterns to be presented, but he determines the types of patterns that the system should be able to process. In other words, the designer, just as in the case of a supervised model, makes a preclassification of the world in terms of what is meaningful to the system. In the case of the "phonetic typewriter" speech samples were selected to cover the space of possible phonemes, and then they were appropriately pre-processed. Clearly, non-supervised learning is an important step in the right direction because, within the preselected set of patterns, the system finds its own categories.

4.3 An attractive formalism

Because of their desirable properties (robustness, learning and generalization capacity), connectionist models or neural networks, are excellent candidates for modeling the interaction of systems with the real world. In fact, they have been widely used for signal processing, pattern recognition, and motor control. The purpose of our critical analysis was to demonstrate that connectionism per so does not automatically resolve the fundamental problems underlying the design of intelligent systems. For example, it does not explain why an agent categorizes the world in the first place, how it focuses on certain aspects of the sensory stimulation and not on others, how it chooses its behaviors depending on the situation, etc. To answer these questions, more is required. And this is what the next section is about.

5 Understanding behavior: design principles of autonomous agents

5.1 Conceptualizing intelligent agents

As pointed out initially, in symbol processing AI high-level capacities like logic, abstract problem solving, human natural language, theorem proving, reasoning and formal games were considered to be the hallmark of intelligence. There was an implicit underlying belief that once we understand the high-level processes we merely add sensors and effectors and we have a system capable of interacting with the real world. Unfortunately, this turned out not to be the case. Doubts have been raised whether this approach of focusing on high-level processes and adding sensors and effectors later on, might not be fundamentally flawed (e.g. Brooks, 1991). Brooks suggested that if we are to understand behavior we must study physically embodied real world agents. This suggestion has lead to a research area which is rapidly growing, namely "New AI".

Researchers in AI realized that what was considered hard initially, turned out to be easy and those things which were viewed as being easy add-ons later like perceptual and motor capabilities, turned out to be hard. Connectionism is an interesting development in this respect since it started focusing on more "low-level" processes such as pattern recognition. But the interest has not only shifted to perceptual-motor skills, but to complete agents. There are a number of reasons for this.

Let us look at a system behaving in the real world, e.g. a mobile robot which has the task to collect uranium ore in an unknown environment. In order to do so it has, among many other things, to avoid obstacles and recognize uranium ore. As it is moving its sensors receive continuously changing physical stimulation and this stimulation is largely determined by what the agent currently does. And what the agent currently does in turn determines, together with the sensory stimulation and the internal state, what it will do next. There is nobody to tell the agent what the relevant patterns of sensory activation are. Unlike supervised and non-supervised neural networks, there is no neat set of training patterns: the agent has to decide from its very own, situated perspective. This constitutes an entirely different set of problems. The design principles have been devised in order to conceptualize agents behaving in a real, physical world.

5.2 *Design principles of autonomous agents*

Types of explanations

Remember that our ultimate goal is to understand principles of intelligence. There is a kind of "meta principle" that has to be endorsed if the design principles are to make sense. It states that agents always be evaluated from three different perspectives, namely functional, learning and development, and evolutionary. Experience has shown that these three perspectives contribute in complementary ways to our understanding.

The *functional* perspective[4] explains why a particular behavior is displayed by an agent based on its current internal and sensory state. Often, this kind of explanation is used in engineering. But also in cognitive science it is highly productive. Just remember Simon's ant on the beach, where it is surprising how seemingly complex kinds of behavior result from very simple mechanisms. The *learning and developmental* perspectives not only resort to internal state, but to some events in the past in order to explain the current behavior. They provide an explanation of how the actual behavior came about. The distinction between learning and development is that development includes maturation of the organism, whereas learning is more general and does not necessarily include change of the organism. *Evolutionary* explanations provide reasons why a particular capacity of an agent is there in the first place, e.g. why it might be beneficial to have a vision system. All these types of explanations can be applied to individuals, but also to groups or whole societies of individuals.

Classical symbol processing models have mostly provided explanations at the functional level. The problem with classical models was, that often no clear distinction was made between internal mechanisms and behavior, because there simply was no behaving organism. Machine learning, in particular connectionist models, have adopted a learning perspective. But the explanations were often relatively uninteresting because the training patterns were prepared by the designer. The evolutionary perspective is a more recent development in AI.

Classes of principles

There are three classes of design principles. An overview is given in Table 1. The first class concerns the kinds of agents and behaviors that are of interest from a cognitive science perspective. We stress the cognitive science perspective since from an engineering perspective, other types of agents are typically of greater interest (at least at the moment). The second class concerns the agent itself, its morphology, its sensors and effectors, its control architecture, and its internal mechanisms. The third class contains principles that have to do with ways of thinking and proceeding, with stances, attitudes, and strategies to be adopted in the design process. Because of space limitations we will only illustrate some of them with a few case studies (for more detail, see Pfeifer, 1996b).

Table 1: Summary of design principles

Principle	Name
Types of agents of interest, ecological niche and tasks	
1	The "complete agents" principle
2	The "ecological niche" principle
Morphology, architecture, mechanism	
3	The principle of parallel, loosely coupled processes (the "anti - homunculus" principle)
4	The "value" principle
5	The principle of sensory-motor coordination
6	The principle of "ecological balance"
7	The principle of "cheap designs"
Strategies, heuristics, stances, metaphors	
8	"Frane-of-reference" principle
9	"Constraints" principles
10	Compliance with principles
	etc.

[4]The term "functional" is used in different ways. Here the term is used to distinguish one level of explanation from a learning/developmental and an evolutionary one.

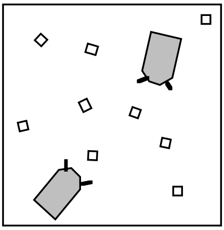
Type of agents, ecological niche, and tasks

Initially we argued that it may not be a good idea to define what we mean by intelligence and that it is perhaps better to define the kinds of agents and behaviors that we are interested in and then look for the mechanisms. This class of principles tries to characterize the kinds of agents that are most worthwhile being investigated. As pointed out before, in classical AI the tasks of interest pertain to high-level thinking. In connectionism they were more in the areas requiring pattern processing, i.e. sensory-motor based. We saw that the study of these tasks alone, still does not suffice to understand intelligence. Thus, this class of principles states that the agents of interest are autonomous (i.e. independent of external control), self-sufficient (i.e. can sustain themselves over extended periods of time), embodied (i.e. they are realized as a physical system), and situated (i.e. the interaction with the environment must be controlled by the agent itself). Moreover, the agent must be able to bring in its own experience in dealing with the current situation. This illustrates the shift of interest in research topics mentioned at the beginning. The basic idea of this set of principles is that if we are to make progress in the study of intelligence, it is these kinds of systems that we must study.

Keep in mind that these design principles only make sense if your primary interest is in cognitive science, i.e. in understanding the basic principles of intelligence. If the main goal is engineering, different principles have to be applied.

5.3 Functional explanations

The example that we will discuss illustrates functional explanations and a principle from class three, namely the "frame-of-reference" principle. Earlier, we discussed the "frame-of-reference" problem using Simon's ant on the beach. The design principle states that in designing and building agents, we have to take the "frame-of-reference" principle into account very carefully. The case study involves a number of simple self-built autonomous robots, the Didabots (Maris and Schaad, 1995). There is an arena with a number of styropor cubes and some Didabots (figure 6). The Didabots are programmed as simple Braitenberg vehicles with only one type of sensor for proximity. All they can do is avoid obstacles. Now look at the sequence of pictures shown in figure 7.



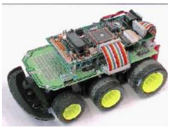


Figure 6: Arena with Didabots and styropor cubes. On the right, a picture of a Didabot is shown.

Initially the cubes are randomly distributed. Over time a number of clusters are forming. At the end there are only two clusters and a number of cubes along the walls of the arena. What would you say the robots are doing?

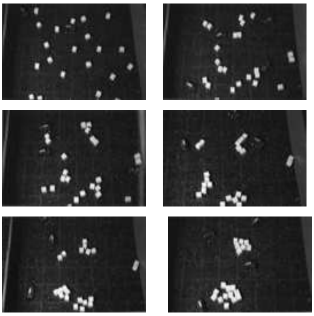


Figure 7: Sequence of situations after a few minutes each. The whole process until a relatively stable situation is achieved lasts roughly 20m.

"They are cleaning up". "They are trying to get the cubes into clusters". These are answers that we often hear. They are fine if we are aware of the fact that they represent an observer's perspective. They describe the behavior. The second answer is a bit problematic since it attributes an intention by using the word "trying". Because we are the designers, we can say very clearly what the robots were programmed to do: to avoid obstacles!

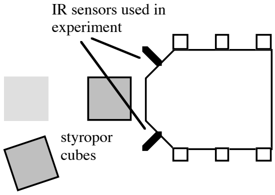


Figure 8: Explanation of the "cleaning up" behavior of the Didabots.

The Didabots only use the two sensors which are marked in black, namely front left and front right. Normally they move forward. If they get near an obstacle within reach of one of the sensors (about 20cm) they simply turn toward the other side. If they encounter a cube head on, neither the left nor the right sensor measure any reflection and the Didabot simply continues moving forward. At the same time it pushes the cube. But it pushes the cube because it doesn't "see" it, not because it was programmed to push it.

For how long does it push the cube? Until the cube either moves to the side and the Didabot loses it, or until it encounters another cube to the left or the right. It then turns away, thus leaving both cubes together (figure 8). Now there are already two cubes together, and the chance that another cube will be deposited near them is increased. Thus, the robots have changed their environment which in turn influences their behavior.

While it is not possible to predict exactly where the clusters will be formed, we can predict with high certainty that one or two clusters will be formed in environments with the geometrical proportions used in the experiment (systematic experiments have been reported by Maris and te Boekhorst, submitted). Thus, we can make predictions, but they are of a different nature than what we are normally used to, say, in physics.

The kind of phenomenon that we have seen in this experiment is also called self-organization. The behavior of the individual leads to a global change, namely the arrangement of the cubes, which in turn influences the behavior of the individuals. As is well-known, self-organization is a ubiquitous phenomenon in the world around us, in biological, social, economic, engineering, and inanimate physical systems.

Because this phenomenon of cluster formation is stable and consistently occurs, given the right conditions and proportions, we can in fact design a cleaning robot without programming explicitly a representation of cleaning into the robots. The task of cleaning is in our minds as designers, it does not have to be in the "minds" of the robots. To use another buzzword, the behavior of such an agent is sometimes called emergent, and the engineering principle "designing for emergence" (Steels, 1991).

5.4 Developmental explanations

Learning and development—a frame-of-reference issue

Developmental explanations (explanations in terms of learning) refer not only to the current situation, but to events in the past. Again, there is a frame-of-reference issue, here. Think about learning, for a moment. What do we mean by the term? Assume that there is an environment with small and large pegs. Initially, the robot will try to pick up all the pegs. After some time the robot only picks up the small ones and ignores the large ones. We call this the current behavior. We then say that the robot has *learned* a distinction between small and large pegs. We can explain the current behavior, for example, by saying that the robot has encountered small and large pegs along the way and found that the large ones are too heavy to pick up. We resort to the history of the interaction of the agent with its environment, in other words, to its individual "experience". But we are *not* saying anything about the internal mechanism. This is not necessary to define learning.

We will proceed by reviewing a number of design principles. They can be applied to all three types of explanations, functional, developmental, and evolutionary. We will first focus on the developmental aspects. In subsection 5.5 we will briefly illustrate some evolutionary considerations.

The "value" principle

This principle states that the agent has to be embedded in a value system, and that it must be based on self-supervised learning mechanisms employing principles of self-organization. If the agent is to be autonomous and situated it has to have a means to judge what is good for it and what isn't. This is achieved by a value system, a fundamental aspect of every autonomous agent.

There is an implicit and an explicit aspect of the value system. In a sense, the whole set-up of the agent constitutes value: the designer decides that it is good for the agent to have a certain kind of locomotion (e.g. wheels), certain sensors (e.g. IR sensors), certain reflexes (e.g. turn away from objects), certain learning mechanisms (e.g. selectionist learning), etc. These values are implicit. They are not represented explicitly in the system. To illustrate the point, let us look at reflexes for a moment. Assume that a garbage collecting robot has the task to collect only small pegs and not large ones. Moreover, it should learn this distinction from its own perspective. The agent is equipped with a number of reflexes: turning away from objects, turning towards an object, and grasping if there has been lateral sensory stimulation over a certain period of time. The value of the first reflex is that the agent should not get damaged. The second and the third reflex increase the probability of an interesting interaction. Note that this interpretation in terms of value is only in the eye of the designer—the agent will simply execute the reflexes.

These reflexes introduce a bias. The purpose of this bias is to speed up the learning process because learning only takes place if a behavior is successful. If the behavior is successful, i.e. if the agent manages to pick up a peg, a value signal is generated. In this case, an *explicit* value system is required. In this way, the intuition that grasping is considered rewarding in itself, can be modeled.

According to the "value" principle, the learning mechanisms have to be based on principles of self-organization, since the categories to be formed are not known to the agent beforehand. Examples are competitive schemes (e.g. Kohonen, 1988a; Martinetz, 1994), or selectionist ones (Edelman, 1987).

This view of value systems and self-organization contrasts with classical thinking. The metaphor of information processing that underlies traditional AI and cognitive science, cannot accommodate self-organization. The "value" principle is supported by many references (e.g. Edelman, 1987; Pfeifer and Verschure, 1992; Pfeifer and Scheier, in press; Thelen and Smith, 1994). It is closely related to the principle of sensory-motor coordination and ecological balance.

The principle of sensory-motor coordination

This principle states that the interaction with the environment is to be conceived as a sensory-motor coordination. Sensory-motor coordination involves the sensors, the control architecture, the effectors, and the agent as a whole. A consequence of this principle is that classification, perception, and memory should be viewed as sensory-motor coordinations rather than as individual modules (e.g. Dewey, 1896; Douglas, 1993; Edelman, 1987).

One of the fundamental problems of visual perception is object invariance. One and the same object—a peg in the case of our robot—leads to large variations in (proximal) sensory stimulation: the latter strongly depends on distance, orientation, lighting conditions, etc. Normally, perception is viewed as a process of mapping a proximal (sensory) stimulus onto some kind of internal representation. The enormous difficulties of classical computer vision to come to grips with the problem of invariances suggests that there may be some fundamental problems involved. Viewing perception as sensory-motor coordination has a number of important consequences.

From an information theoretic view, the sensory-motor coordination leads to a dimensionality reduction of the high-dimensional sensory-motor space (Pfeifer and Scheier, in press). This reduction allows learning to take place even if the agent moves. In fact, movement itself is beneficial since through its own movement, the agent *generates* correlations in the interaction with the environment. The second important aspect of sensory-motor coordination is the generation of cross-modal associations, including proprioceptive cues originating from the motor system (Thelen and Smith, 1994; Scheier and Lambrinos, 1996).

Figure 9: Infant categorizing objects and building up concepts while engaged in sensory-motor coordination.

Additional support for the principle of sensory-motor coordination comes from developmental studies. There is a lot of evidence that concept formation in human infants is directly based on sensory-motor coordination (Thelen and Smith, 1994; Smith and Thelen, 1993; see figure 9). The concepts of humans are thus automatically "grounded". Similarly, if this principle is applied to artificial agents, the latter will only form fully grounded categories. The symbol grounding problem is really not an issue—anything the agent does will be grounded in its sensory-motor coordination. Note that the terms categorization and concept building are entirely observer-based. They relate only to the behavior of the infant, not to any sort of internal mechanism.

There is another kind of approach that closely relates to this principle, namely active vision (e.g. Ballard, 1991). Vision is not seen as something that concerns only input, but movement is considered to be an integral aspect.

As already alluded to, this view contrasts with the traditional view of perception as a process of mapping a proximal stimulus onto an internal representation. In the view proposed here, the object representation is in the sensory-motor coordination. "Recognizing" an object implies re-enacting a sensory-motor coordination. Most objections to this view of perception have their basis in introspection. The latter has long ago been demonstrated to be a poor guide to research (Nisbett and Wilson, 1977). This principle is supported by numerous research contributions (e.g. Ballard, 1991; Dewey, 1896; Douglas, 1993; Edelman, 1987; Thelen and Smith, 1994; Smith and Thelen, 1993; Scheier and Lambrinos, 1996; Pfeifer and Scheier, in press; Scheier and Pfeifer, 1995)

The principle of "ecological balance"

The principle of "ecological balance" states that there has to be a match between the "complexity" of the sensors, the actuators, and the neural substrate. Moreover, it states that the tasks have to be "ecologically" adequate. The way the term "complexity" is used here, appeals to our everyday understanding: a human hand is more complex than a forklift, a CCD camera more complex than an IR sensor.

From this principle we can get considerable leverage. Let us look at an example illustrating how *not* to proceed. Assume that we have a robot with two motors and a few IR sensors, say the robot Khepera™. In some sense, this design is balanced due to the intuition of the engineers that built it (except that its processor is too powerful if it is fully exploited). Assume further that some researchers have become frustrated because with the IRs they can only do very simple experiments. They would like to do more interesting things like landmark navigation.

The next logical step for them is to add a CCD-camera. It has many more dimensions than the few IR sensors. The rich information from the camera is transmitted to a central device where it is processed. This processing can, for example, consist in extracting categories. But the categories are formed as a consequence of a sensory-motor coordination. Because the motor system of the agent is still the same, the resulting categories will not be much more interesting than before (although they may be somewhat different). Trying to build categories using only the visual stimulation from the camera (not as a sensory-motor coordination) would violate the principle of sensory-motor coordination. Classical computer vision has violated this principle—and the problems are well-known. It would be a different story if, together with the CCD camera, additional motor capabilities would have been added to the robot, like a gripper or an arm of sorts. Figure 10 shows a balanced design on the left, an unbalanced one in the middle, and again a more balanced one on the right.

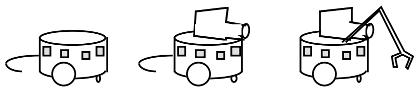


Figure 10: Balanced design on the left, unbalanced design in the middle, and again more balanced design on the right.

An approach that is fully compatible with the principle of "ecological balance" is the Cog project for developing a humanoid robot (Brooks, 1994; Brooks and Stein, 1993). More sophistication on the sensor side (two eyes, each with a camera for peripheral and foveal vision), is balanced by more complexity on the motor side. The arm and the hand are quite sophisticated. Moreover, the head and the eyes can all move which leads to a system of a very large number of degrees of freedom. A lot of the processing is done peripherally, and the central processing capacity is not inflated artificially. It is not surprising that Cog fulfills this design principle. It was Brooks who pointed out that tasks need to be ecologically appropriate (Brooks, 1990). In particular he argued that "elephants don't play chess." We couldn't agree more.

Important evidence for this principle comes also from studies in infant psychology by Bushnell and Boudreau (1993). Their results suggests that there is in fact a kind of co-evolution in the sensory-motor development of the infant. Roughly speaking, acuity of visual distinctions highly correlates with precision of motor movement.

Again, this view sharply contrasts with traditional AI and cognitive science, where intelligence was seen as centralized information processing, with no, or very little consideration given to the physical set-up. A concept like "ecological balance" would not make sense in that framework. References supporting this principle include Brooks, 1991, 1994; Pfeifer, 1995; Smith and Thelen, 1993; Bushnell and Boudreau, 1993.

The principle of parallel, loosely coupled processes

Let us just mention very briefly this principle without going into any detail. In essence, it states that intelligence is emergent from a large number of parallel, loosely coupled processes. These processes run asynchronously and are largely peripheral, requiring little or no centralized resources. It could also be called the "anti-homunculus" principle. One of the main claims here is that coherent behavior can be achieved without central control. A beautiful example that fully endorses this principle is, again, the Cog project (Brooks, 1994; Brooks and Stein, 1993). In our own work we have applied this principle to all our agents (e.g. Scheier and Pfeifer, 1995; Pfeifer and Scheier, in press).

This principle contrasts sharply with classical thinking where a centralized seat of intelligence is assumed. Classical thinking does not object to parallel processes (as we have seen in connectionism). The objection is that coherence cannot be achieved unless there is central integration. The classical view that maintains integration is necessary is especially predominant in psychology, in particular cognitive psychology.

A lot of research in the field of cognitive science and "New AI" supports the principle of parallel, loosely coupled processes (e.g. Braitenberg, 1984; Brooks, 1991; Brooks and Stein, 1993; Maes, 1991; Steels, 1992; Scheier and Pfeifer, 1995; Pfeifer and Scheier, in press).

The principle of "cheap designs"

The last principle that we will discuss is the one of "cheap design". It states that good designs are "cheap". "Cheap", as used here, has several meanings. For the purposes of this paper, it means parsimonious. Moreover, this parsimony is to be achieved through exploitation of the physics. A nice illustration is insect walking. Leg coordination in insects does not require a central controller. There is no internal process corresponding to global communication between the legs, they communicate only locally with each other (e.g. Cruse, 1991). But there *is* global communication between all the legs, namely through the environment. It is mediated by a physical process, not by an information process (or a process of signal transfer) within the agent. If the insect lifts one leg, the force on all other legs is changed instantaneously because of the weight of the insect. This saves the insect a lot of—unnecessary—neural substrate. Another example of a cheap design are the Didabots which are cleaning up the arena of styropor cubes. There is ample evidence supporting this principle (e.g. e Brooks, 1991; Cruse, 1991; Horsewill, 1992; Franceschini et al., 1992; Pfeifer, 1993, 1995; Thorpe and Imbert, 1989.

5.5 Evolutionary explanations

We do not want to overstress this point, but some of the design principles have interesting evolutionary interpretations. Take, for example, the principle of "ecological balance". Natural designs are ecologically balanced. It seems that evolution favors balanced designs. Recent developments in evolutionary robotics have suggested simulated evolution as a design principle (e.g. Harvey et al., in press). It would be interesting to see whether eventually balanced designs will emerge from these efforts. Of course, this would require evolving complete agents, not only control architectures (as is currently done).

In our research we have mostly been focusing on the functional and the learning/developmental perspectives. In the future we will include evolutionary principles into our considerations.

6 Discussion

We have now completed our argument. We began by pointing out some fundamental problems of symbol processing models and defined the "information processing view". We then showed to what extent connectionist models resolve some of these issues. They represent an important development in the right direction, because they process patterns of activation rather than symbols. But we saw that connectionist models—for the better part—remain within the information processing paradigm. They typically are based on basic (symbolic) categories that are defined by the designer. This holds for supervised as well as for non-supervised schemes. A very different kind of thinking is needed if we are to understand and design systems which have to interact with the real world. And it seems, that the history of AI, cognitive science, and robotics has taught us that intelligence always requires the interaction with the real physical world.

The paradigm of “New AI”, of employing embodied physical agents (typically in the form of autonomous robots) as a research tool, helps us asking the right, questions. The questions that we have to ask, relate to behavior of complete agents. It is amazing how much our view of intelligence changes with this perspective. All of a sudden, it seems possible to overcome some of the fundamental problems of the traditional approach.

An example is the symbol grounding problem. As we have seen, the categories and concepts an agent acquires, will be grounded if we focus on sensory-motor coordination. We have not “resolved” the symbol grounding problem because it does not need to solved. But we have shown how we can design agents without getting trapped in it. Likewise, it will not be possible to entirely eliminate the frame problem. However, the principle of "ecological balance" tells us that we should not artificially increase the complexity of the neural substrate (which would be necessary if the agent were to build sophisticated models of the environment) if the sensory-motor system remains the same. Thus, if we observe this design principle, we are much less in danger of building models that are too complex for the specific agent-environment interaction. Or recall the computational problems involved in perception. By viewing perception as sensory-motor coordination, the computational complexity can be dramatically reduced. Note that as a side-effect of applying the design principles, real-time performance increases because less processing is required.

There is another point of concern. Initially, when discussing the classical approach we introduced the notion of representation. There are these mappings between the outside world and the agent (called “encode” and “decode”). Connectionst models typically do not deal appropriately with these mappings, because they are given by the input and output categories of the model (letters and phoneme features). They are predefined by the designer, not acquired by the model itself. In real-world agents, this mapping is mediated by the physics of the agent. It turns out that if we are trying to interpret the weight patterns and activation patterns in a neural network, this is only possible if we know how the sensory and motor systems function, and *where they are physically positioned on the robot*. Trying to find where categories are represented—remember that the categories are observer-based (frame-of-reference)—in a network, is a task that can only be achieved if it is exactly known how this network is embedded in the agent. Otherwise, activation levels and connection strengths have *no meaning*: they cannot be abstracted out. Thus, representation is no longer a property of a formalism, or of a mapping, but a property of a complete agent.

As a last point for discussion let me anticipate your questions, namely “will this approach of ‘New AI’ ever scale up? Will we ever be able to solve the same kinds of problems that we have been able to solve using the classical approach?" The answer depends on what we mean by the sentence “that we have been able to solve using the classical approach”. It could mean that we have been able to build systems that support humans in their work. In this case, the research goal has not been to develop *intelligent* systems but useful computer systems. If the intended meaning is that we have been able to model human expertise, I would argue that—perhaps with some exceptions—classical models only capture very limited aspects of human intelligence (or expertise), and perhaps not the most interesting ones.

And this brings me to the concluding remark. It has often been suggested that for the low-level competences, the sensory-motor aspects, connectionist models, or “New AI” style models might be appropriate, but that for the high-level part we may need to resort to symbol processing concepts. While from an engineering perspective there is little to be argued if it works, from a cognitive science perspective, it can be predicted that this approach will not lead to interesting insights. The reason is that this way of proceeding constitutes a category error: on the one hand the categories are built up by the agent itself, i.e. they are agent-based, whereas the ones used in the symbolic system, are designer-based—once again, a "frame-of-reference" issue.

The design principles outlined above do not cover all the insights of the very rich field of "New AI". But we do believe that they capture a large part of the most essential aspects of what has emerged from pertinent research. The principles described may seem somewhat vague and overly general, but they are enormously powerful as heuristics, providing guidelines as to what sorts of experiments to conduct next and what agents to design for future experiment. In order to achieve some degree of generality we have deliberately left out a lot of detail. These principles not only help us evaluate existing designs, but they get us to ask the right questions.

In the future we might be looking for something more formal, than merely a set of verbally stated design principles. Eventually, this will certainly be necessary. But what this "theory" will look like, is entirely open.

What is needed right now is an in-depth discussion of these design principles. They have to be revised and the list of principles has to be augmented.

7 Conclusions

I hope that it has been shown, that if we are to understand intelligence, we need more than the classical tools of symbol processing AI. But connectionism alone will not solve the fundamental problems either. We need to take the physics of the agent and how it interacts with its environment into account. In other words we need to go beyond the information processing metaphor. We have outlined a number of principles that will hopefully form the basis for a productive discussion.

Acknowledgments

This paper includes materials from various sources. The "Encyclopedia of Microcomputers" (Pfeifer, 1996a), the "Proceedings of the 4th International Conference on the Simulation of Behavior: From Animals to Animats" (Pfeifer, 1996b), and the "Journal of Robotics and Autonomous Systems" (Pfeifer, 1995). It also contains material from a presentation given at ER'96 at the Canadian Embassy in Tokyo, April 1996. The research reported was performed largely under the auspices of the Swiss National Science Foundation, grants # 20-40581-94 and 32-37819.93. I am grateful to Fumio Mizoguchi and Ryuichi Oka for inviting me to write this paper. I would like to thank Christian Scheier, Dimitri Lambrinos, and Ralf Salomon for their inspiring discussions and valuable comments on the manuscript.

References

Ballard, D.H. (1991). Animate vision. *Artificial Intelligence*, **48**, 57-86.

Beer, R. (in press). The dynamics of adaptive behavior: a research program. To appear in: *Robotics and Autonomous Systems, Special Issue on "Practice and Future of Autonomous Agents",* R. Pfeifer, and R. Brooks (eds.).

Braitenberg, V. (1984). *Vehicles: experiments in synthetic psychology*. Cambridge, Mass.: MIT Press.

Brooks, R.A. (1990). Elephants don't play chess. *Robotics and Autonomous Systems*, **6,** 3-15.

Brooks, R.A. (1991). Intelligence without representation. *Artificial Intelligence*, **47**, 139-160.

Brooks, R.A. (1994). Coherent behavior from many adaptive processes. In: D. Cliff, P. Husbands, J.-A. Meyer, and S.W. Wilson (eds.). *From animals to animats 3. Proc. SAB'94,* 22-29.

Brooks, R.A., and Stein, L.A. (1993). Building brains for bodies. Memo 1439, MIT Artificial Intelligence Laboratory, Cambridge, Mass.

Bushnell, E.M. and Boudreau, J.P. (1993). Motor development in the mind: The potential role of motor abilities as a determinant of aspects of perceptual development. *Child Development*, **64**, 1005-1021.

Chapman, D. (1987). Planning for conjunctive goals. *Artificial Intelligence*, **32**, 333-337.

Clancey, W.J. (1991). The frame of reference problem in the design of intelligent machines. In K. van Lehn (ed.). *Architectures for intelligence*. Hillsdale, N.J.: Erlbaum.

Cognitive Science (1993), **17**. Norwood, N.J.: Ablex Publ.

Cruse, H. (1991). Coordination of leg movement in walking animals. In: J.-A. Meyer, and S.W. Wilson (eds.). *From animals to animat 1. Proc. SAB'90,* 105-119.

Dennett, D. (1987). Cognitive wheels. In Z.W. Pylyshyn (ed.) (1987). *The robot's dilemma. The frame problem in artificial intelligence*. Norwood, N.J.: Ablex (2nd printing 1988).

Dewey, J. (1896). The reflex arc concept in psychology. *Psychol. Rev.*, **3** (1981) 357-370; Reprinted in: J.J. McDermott (ed.) *The Philosophy of John Dewey*. Chicago, IL: University of Chicago Press, 136-148.

Douglas, R.J., Martin, K.A.C., and Nelson, J.C. (1993). The neurobiology of primate vision. *Bailliere's Clinical Neurology*, **2**, No. 2, 191 - 225.

Edelman, G.E. (1987). *Neural Darwinism. The theory of neuronal group selection.* New York: Basic Books.

Franceschini, N., Pichon, J.M., and Blanes, C. (1992). From insect vision to robot vision. *Phil. Trans. R. Soc. Lond. B*, **337**, 283-294.

Harnad, S. (1990). The symbol grounding problem. *Physica D*, **42**(1-3), 335-346.

Harvey, I., Husbands, P., Cliff, D., Thompson, A., Jakobi, N. (1996). Evolutionary robotics: the Sussex approach. *Robotics and Autonomous Systems, Special Issue on "Practice and Future of Autonomous Agents"*, R. Pfeifer, and R. Brooks (eds.).

Horsewill, I. (1992). A simple, cheap, and robust visual navigation system. In: J.-A. Meyer, H.L. Roitblat, and S.W. Wilson (eds.). *From animals to animats 2. Proc. SAB'92,* 129-137.

Janlert, L.E. (1987). Modeling change—the frame problem. In Z.W. Pylyshyn (ed.) (1987). *The robot's dilemma. The frame problem in artificial intelligence.* Norwood, N.J.: Ablex (2nd printing 1988).

Kohonen, T. (1988a). *Self-organization and associative memory*. Berlin: Springer.

Kohonen, T. (1988b). The "Neural" phonetic typewriter. *IEEE Computer*, 1988, March, 11-22.

Laird, J.E., Newell, A., and Rosenbloom, P.S. (1987). SOAR: an architecture for general intelligence. *Artificial Intelligence,* **33**, 1-64.

Maes, P. (1991). A bottom-up mechanism for behavior selection in an artificial creature. In: J.-A. Meyer, and S.W. Wilson (eds.). *From animals to animats 1. Proc. SAB'90,* 238-246.

Maris, M., and Schaad, R. (1995). The didactic robots. University of Zurich, AI Lab Techeport, #95.05.

Maris, M., and te Boekhorst, R. (submitted). Exploiting physical constraints: heap formation through behavioral error in a group of robots (submitted to IROS'96).

Martinetz, T. (1994). Topology representing networks. *Neural Networks*, **7**, 505-522.

McCarthy, J., and Hayes, P.J. (1969). Some philosophical problems from the standpoint of artificial intelligence. In: B. Meltzer, and D. Michie (eds.), *Machine Intelligence,* **4**, 463-502.

McFarland, D., and Bösser, M. (1993). *Intelligent behavior in animals and robots*. MIT Press.

Nisbett, and Wilson (1977). Telling more than we can know: verbal reports of mental processes. *Psychological Review*, **84,** 231-259.

Newell, A. (1990). Unified theories of cognition. Cambridge, Mass.: Harvard University Press.

Newell, A., and Simon, H.A. (1976). Computer science as empirical inquiry: symbols and search. *Comm. of the ACM*, **19**, 113-126.

Pfeifer, R. (1993). Cheap designs: exploiting the dynamics of the system-environment interaction. Three case studies on navigation. In: Conference on *"Prerational Intelligence"*. Center for Interdisciplinary Research, University of Bielefeld, 81-91.

Pfeifer, R. (1995). Cognition—perspectives from autonomous agents. *Robotics and Autonomous Systems*, **15**, 47-70.

Pfeifer, R. (1996a). Symbols, patterns, and behavior: beyond the information-processing metaphor. In A. Kent and J.G. Williams (eds.). *Encyclopedia of Microcomputers*, Vol. **17**, 253-275. New York: Marcel Dekker, Inc.

Pfeifer, R. (1996b). Building "Fungus Eaters": Design principles of autonomous agents. *Proc. SAB'96.*

Pfeifer, R., and Scheier, C. (in press). Sensory-motor coordination: the metaphor and beyond. To appear in: *Robotics and Autonomous Systems, Special Issue on "Practice and Future of Autonomous Agents"*, R. Pfeifer, and R. Brooks (eds.).

Pfeifer, R., and Verschure, P.F.M.J. (1992). Distributed adaptive control: a paradigm for designing autonomous agents, in F.J. Varela, and P. Bourgine (eds.) *Proc. ECAL-92*, 21-30.

Putnam, H. (1975). Philosophy and our mental life. In H. Putnam (ed.) *Mind, language and reality: philosophical papers, vol. 2*. Cambridge: Cambridge University Press.

Pylyshyn, Z.W. (ed.) (1987). *The robot's dilemma. The frame problem in artificial intelligence*. Norwood, N.J.: Ablex (2nd printing 1988).

Quillian, R. (1968). Semantic memory. In M. Minsky (ed.). *Semantic information processing*. Cambridge, Mass.: MIT Press.

Reeke, G.N. Jr., and Edelman, G.M. (1988). Real brains and artificial intelligence. *Daedalus*, **117**(1), 143-173.

Rumelhart, D., and McClelland J. (1986). *Parallel distributed processing*. Cambridge, Mass.: MIT Press.

Scheier, C., and Lambrinos, D. (1996). Categorization in a real-world agent using haptic exploration and active vision. *Proc. SAB'96.*

Scheier, C., and Pfeifer, R. (1995). Classification as sensory-motor coordination: a case study on autonomous agents. *Proc. ECAL-95,* 657-667.

Sejnowski, T.J., and Rosenberg, C.R. (1987). Parallel networks that learn to pronounce English text. *Complex Systems*, **1**, 145-168.

Simon, H.A. (1969). *The sciences of the artificial.* Cambridge, Mass.: MIT Press.

Smith, L.B., and Thelen, E. (eds.) (1993). *A dynamic systems approach to development. Applications.* Cambridge, Mass.: MIT Press, Bradford Books.

Steels, L. (1991). Towards a theory of emergent functionality. In: J.-A. Meyer, and S.W. Wilson (eds.). *From animals to animats 1. Proc. SAB'90,* 451-461.

Steels, L. (1992). The PDL reference manual. VUB AI Lab memo 92-5.

Steinhage, A., and Schöner, G. (in press). Self-calibration based on invariant view recognition: Dynamic approach to navigation. To appear in: *Robotics and Autonomous Systems, Special Issue on "Practice and Future of Autonomous Agents",* R. Pfeifer, and R. Brooks (eds.).

Suchman, L. (1987). *Plans and situated action.* Cambridge University Press.

Thelen, E. and Smith, L. (1994). *A dynamic systems approach to the development of cognition and action.* Cambridge, Mass.: MIT Press, Bradford Books.

Thorpe, S.J., and Imbert, M. (1989). Biological constraints on connectionist modelling. In R. Pfeifer, Z. Schreter, F. Fogelman-Soulié, and L. Steels (eds.). *Connectionism in Perspective.* Amsterdam: North-Holland, 63-92.

Verschure, P.F.M.J. (1992). Taking connectionism seriously: the vague promise of subsymbolism and an alternative. *Proceedings of The Fourteenth Annual Conference of The Cognitive Science Society.*

Winograd, T. and Flores, F. (1986). *Understanding computers and cognition.* Reading, Mass.: Addison-Wesley.

Creative AI Lab
creative-ai.org

Design by *doubledate*