Synthesis Lectures on Computer Science

SYNTHESIS
COLLECTION OF TECHNOLOGY

Norman I. Badler

On Raising a Digital Human

A Personal Evolution

Springer

Synthesis Lectures on Computer Science

The series publishes short books on general computer science topics that will appeal to advanced students, researchers, and practitioners in a variety of areas within computer science.

Norman I. Badler

On Raising a Digital Human

A Personal Evolution

Springer

Norman I. Badler
Cesium GS and The University of Pennsylvania
Philadelphia, PA, USA

ISSN 1932-1228 ISSN 1932-1686 (electronic)
Synthesis Lectures on Computer Science
ISBN 978-3-031-63944-9 ISBN 978-3-031-63945-6 (eBook)
https://doi.org/10.1007/978-3-031-63945-6

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To my colleagues, students, postdocs, staff, and family who have made this story possible, exciting, and meaningful

Foreword

Modeling virtual humans is a critical capability for engineers and designers to incorporate real-time interactive manipulation and display of human figures into computer-aided design systems. For design engineers, such models can also inform on how to control and design human-centered systems central to the emergent concept of concurrent engineering. Additionally, modeling a large group of virtual humans, i.e., computational methods and models for virtual crowd simulation of agents with artificial intelligence (AI), is considered one of the essential components of reproducing and evaluating egress performance in specific scenarios to improve human safety and comfort in built environments as well as for creating believable virtual environments for numerous applications. For example, understanding pedestrian behaviors during egress situations is of considerable importance in such contexts. Evaluation, analysis, and comparisons of crowd simulation data, derived from real-world experiments, are often required for a wide range of architectural and engineering designs in large-scale structures and/or engineered artifacts (e.g., buildings, architectures, bridges, aircraft, trains, public transit systems, etc.) and also for digital production of online educational materials and story-telling.

Dr. Norman (Norm) Badler is one of the earliest Computer Graphics pioneers who led the research in Computer Animation for human modeling and crowd simulation. For nearly five decades, his research has focused on designing algorithms and developing computer software for computer graphics with a focus on *modeling virtual humans.* He is widely known for his extensive work on 3D computer graphics representations of the human body and his expansive research to acquire, simulate, animate, and control human body, face, gesture, locomotion, and manual task motions. These virtual humans are intended to portray physical, cognitive, perceptual, personality, relationship, and cultural parameters and are used singly, in small groups, or in crowd masses. These digital models exhibit physiological, self-initiated (autonomous), task-dependent, reactive, interpersonal, or social behaviors. They are presented through 3D computer graphics systems that allow interactive visual experiences suitable for training or real-time immersion in populated virtual worlds.

Beyond the science and engineering of digital humans and virtual crowds, the personal narrative of this book is a detailed chronicle of how Norm's passion for computer science and mathematics started to evolve into a thriving and distinguished career as the leading authority on virtual human modeling today. Beginning in grade school, Norm was captivated by digital computers and polyhedra, which led him to master programming by high school. He started by exploring and pursuing his interests starting at the University of California, Santa Barbara then the University of Toronto, where he finished his doctoral research before eventually establishing a distinguished academic career at the University of Pennsylvania. Growing up in an entirely modest and non-academic family, his path is a tale of luck, coincidences, circumstantial exploration, and transformative experience. Throughout, he was fortunate to encounter and collaborate with mentors, colleagues, and especially his wife, Virginia, and two sons, Jeremy and David. Along the way, Norm pioneered novel approaches, software systems, and impressive prototypes that pushed the frontier of human modeling and computer animation. He has educated generations of researchers and students, and he bridged computer graphics and other scientific and engineering disciplines, including anthropometry, cognitive science, and ergonomics. Norm's impact can be truly felt across education, entertainment, academia, and industry around the world!

By highlighting Norm's journey from elementary school to an incredibly successful career in computer science as a world-renowned leading authority in Computer Graphics, especially on modeling and animation of virtual humans, this book showcases the importance of pursuing one's passions with determination and hard work. Norm's persistence and curious mind led him to a stellar culmination of incredible achievements ranging from research to teaching, computer vision to computer graphics, digital humans to animated crowds, and articulated movement to expressive emotions and cognitive behaviors. This autobiography of Prof. Norm Badler serves as an inspiration to all of us with its many important lessons on relationships between people and across communities as well as the essential role that real, industrial-strength problems have in driving virtual human research and more.

July 2024

Ming C. Lin
Barry Mersky and Capital One E-Nnovate
Endowed Professor, Distinguished
University Professor
University of Maryland
College Park, MA, USA

Preface

This book tells the story of building digital virtual human models in the context of the background, choices, and occurrences that shaped the author's own involvement and personal evolution. Such digital models found motivating problems and applications in engineering, anthropology, medicine, and crowd simulation, and numerous connections to other disciplines have informed and enriched their design, development, and deployment.

Encouragement to tell this story originated from several sources. Dave Kasik and Mary Whitton asked me if I would like to write a personal essay sometime in 2023 on "The Story of Virtual Humans" for *IEEE Computer Graphics and Applications.* That article, for its "Origins" series, appears in the November/December 2023 issue. It is a much-abridged version of this full account, motivated by Patrick Cozzi. He independently asked if I would do a similarly themed piece as a series of blog posts in my role as Head of Research at Cesium. After I retired in 2021 from my tenured position on the Computer and Information Science faculty at the University of Pennsylvania, I had toyed with the idea of committing memories to text before they were gone forever. Personal reasons further incentivized me to start writing in late 2022. I would not have undertaken such an enterprise today had I not experienced the joy of non-technical writing for the first time when I spent summer 2017 authoring, with my wife Virginia, our book *Dachshund Days.*

I have generously used citations to provide pointers into relevant literature, fully recognize co-authors, and provide external evidence for claims. This millennium has seen an explosion of interest in virtual beings—from both the academic and corporate spheres—which would require immense effort to properly survey. This is not a survey about today, but about how we got here.

For a proper accounting, I had to reach back into grade school in the 1950s. I can then tell the full story of computers, graphics, and virtual humans within a context of events, coincidences, good fortune, and influential people who have contributed to shaping my life and career. I have received good advice and bad. Sometimes the latter, by being ignored, turned out to be the best of all. Perhaps you may even find a few useful life pointers along the way. I hope you will enjoy following my journey.

This account is for anyone who wants to understand the history of virtual digital human beings, how they evolved, and especially how they must address numerous human characteristics to achieve any sense of "human-ness." The discussions are deliberately kept accessible to a modestly technical-savvy audience, avoiding algorithmic details and limiting unnecessary jargon. Extensive references to the literature fill those gaps. It is not an "art" book, either. While the visual appearance of virtual humans can be nearly flawless today, the underlying properties and controls these figures require are the more important aspects of this book. It could be a motivator to explore the development and applications of digital humans today, whether as a user, researcher, teacher, or just as an observer of the amazing likenesses we can see in games, movies, and the Internet. Digital human background history has not yet been adequately coalesced into a single volume. This account tries to remedy that and includes a deep personal and evolutionary perspective to humanize the story.

Haverford, PA, USA Norman I. Badler
2024

Acknowledgments

I have endeavored to recognize the numerous individuals who enabled, participated in, and fundamentally created this story. I hope I have provided fair and accurate credit and assessments and apologize to any who were accidentally omitted. By including extensive references, all contributors should be remembered for their roles. I have also tried to acknowledge, within the stories themselves, the external sponsors who provided the problems, the funding, the talk invitations, and the intellectual and engineering contexts for many of these endeavors. I have also tried to fully describe and express my appreciation for numerous University of Pennsylvania faculty collaborators over the years: Aravind Joshi, Ruzena Bajcsy, Howard Morgan, Stephen Smoliar, Bonnie Webber, Mark Steedman, Martha Palmer, Mitch Marcus, Michael Greenwald, Stephen Lane, Dimitris Metaxas, Ani Nenkova, Barry Silverman, Renata Holod, Alla Safanova, Ladislav Kavan, Chenfanfu Jiang, and Clark Erickson. I realized that I needed to add the motivational Prologue when describing the book to Penn English Prof. Al Filreis. Numerous students, staff, postdocs, and visiting colleagues all deserve thanks for the multiplicity of influences they have originated, authored, and contributed over five decades. I tried to give them their rightful credit here.

The process of writing takes time and focus. Retiring from my tenured faculty position at the University of Pennsylvania provided some time to consider writing. A huge incentive came from colleagues David Kasik and Mary Whitton, who encouraged me to document the *Jack* human modeling software story. “On Raising a Virtual Human” was subsequently published in the November/December 2023 issue of *IEEE Computer Graphics and Applications* (DOI 10.1109/MCG.2023.3320319; 0272-1716 © 2023 IEEE). This material is reprinted here with permission from IEEE.

Another great impetus came from Patrick Cozzi, CEO of the Philadelphia geospatial software company Cesium. After retiring from Penn, Cozzi hired me into a new role as Cesium Head of Research and suggested that I write a blog about my earlier work on digital humans. That idea evolved rapidly into the full story documented here. Cozzi’s support and encouragement were invaluable and much appreciated.

Oddly enough, however, the commitment to write actually came with our adoption of two dachshund puppies, Oscar and Vienna, in late 2022. At the end of 2022 and into January 2023, we had as a houseguest a relative who had to be separated from them. So the dogs and I spent hours sequestered in the family room. I had dedicated time to begin writing.

Throughout my career, and as recounted here, my wife Ginny and our two sons Jeremy and David have been inspirational. I have been especially happy to recount how Ginny was not just supportive but fundamentally motivational and essential to my intellectual evolution. I am convinced my career path would have been quite different—and likely not nearly as satisfying and rewarding—had she not been by my side since our marriage in 1968. Although my parents passed away before this work was completed, I believe I inherited my algorithmic mindset from my father Bernard, and any artistic appreciation from my mother Lillian.

Special thanks are due to those who read, commented on, and corrected portions of earlier drafts. Dave Kasik, Mary Whitton, Bonnie Mitchell, and other anonymous reviewers improved my *IEEE CG&A* Origins versions, and these modifications helped here. I especially note editorial and content contributions from Andie Tursi of Cesium, Amy Calhoun of the University of Pennsylvania, Funda Durupinar of the University of Massachusetts, Bonnie Webber and Mark Steedman of the University of Edinburgh, Aline Normoyle of Bryn Mawr College, my wife Ginny, and my sons Jeremy and David Badler. I appreciate permission to use images from Self Magazine, Lynn Oseguera, and Michael Zyda, and of real people Virginia Badler, Jeffry Nimeroff, and Wang Gu.

In my quest to find a publisher, my long-time friend and prior publisher Mike Morgan proved invaluable by recommending this work to Christine Kiilerich of Springer Nature. I am grateful for her support and the Springer Nature staff in bringing this to fruition.

Contents

Prologue

1

"Who is real?" The question sounds odd, even vaguely ungrammatical. Using "who" implies a reference to a person; anything else fares more sensibly with "what." As we approach the quarter century mark, the question becomes meaningful and even necessary. Just because someone looks real does not necessarily imply that they are. We watch motion picture actors who are replaced by digital doubles, who may no longer be alive, or who are imaginative versions of someone who is portrayed as if they have been alive. We become immersed in the story. The acting and crafting of such re-animations are good enough to visually convince the viewer that they present as "who" and not "what." Already the technologies and skills are accessible enough for the creation of virtual news announcers, guides, weather reporters, and customer service representatives. Marketing is a key sector, as virtual influencers are always available to make a pitch and real money can flow for goods, fashions, and services. Only a few decades ago we experienced the first telephone call centers with synthetic voice-response systems for reservations and straightforward questions. We could detect almost immediately that these interlocutors were not real and that we could often escape to a live voice by pressing "0." Visual depictions of interacting virtual people now confound that simple formula. They seem to act, or at least look, like people, so our relationships with them, whether positive or negative, are influenced by visual, cognitive, and emotional cues honed over millennia of our own survival and evolution.

"Who is real?" begs the broader question of "What makes something real?" The question, and possible answers, have engaged philosophers since philosophy itself existed. Does Descartes' "I think, therefore I am" apply to a virtual *digital* being? If it has the appearance and cognitive abilities of a human, does that make it human? Current discussions debate the role and ethics of using machine learning to construct artificial

N. Badler, *On Raising a Digital Human*, Synthesis Lectures on Computer Science,
https://doi.org/10.1007/978-3-031-63945-6_1

intelligence systems that can compete with human intelligence. The Turing test, involving human-like conversational responses, is no longer a measure of a human intellectual presence. Nor is tangibility essential for "Who is real?" Humanoid robots are coming. They are physical and hence certainly tangible, but technological difficulties still simplify the problem of assessing reality. Something tells us they are not human. The roboticist Masahiro Mori observed that effect in 1970 with his invention of the phrase "the uncanny valley" [1]. As our artifacts come to resemble humans, they reach a point where the forced realism itself is subliminally detectable and visually disturbing. The best way to avoid the uncanny valley is to emphasize some nonhuman attributes of our reproductions. We usually have no "reality" issues with cartoons, monsters, or industrial robots. We can still attribute human qualities to their behaviors without them taking on accurate human forms.

Defining "human-ness" is way beyond my horizon. As a computer scientist, I try to reduce complexities to structures and algorithms. To start, one must decide which human dimensions are most important, contributory, or even relevant. Analysis by narrative is a rich and endless literary route for describing human characters, but it does not provide any essential insights into how to frame "realness" algorithmically. Analysis by decomposition is a possible strategy: by developing models for the various attributes and abilities that make us human, we might be able to design and build computational components that contribute to a more realistic whole. There are no guarantees that even if we succeed in modeling some dimensions, we will be able to reassemble them into a coherent, consistent, and complete being. Entire fields of human scientific endeavor, such as psychology, cognitive science, neuroscience, anthropology, and biomechanics, have attacked some of these component dimensions. As deep and powerful as these explorations are, they historically saw little need to marry their perspectives with those of the other communities. Cross-fertilization and interdisciplinary studies have emerged as viable research directions in the past few decades. Interconnections are challenging, but the re-integration of at least a portion of what is known will be increasingly motivated by emergent applications for digital humans.

Where to start? There are many dimensions to human beings, so the first task is to decide which ones are both meaningful and computationally manageable. One possible organization appears in Fig. 1.1. In addition to the obvious *Visual Realism* aspect, I chose dimensions of *Personal Familiarity* and *Interpersonal History*. I do not think these are commonly addressed, but the three together offer a characterization of many of the present applications for digital beings. Since these categories are intended for real-time roles—as opposed to pre-recorded movie characters, for example—a real person or observer is engaged in some fashion. I will call these roles, rather interchangeably, users, players, observers, or interlocutors. Usually, context informs which term would be most appropriate. The distinctions are not particularly meaningful here.

What do the labels on my dimensions mean?

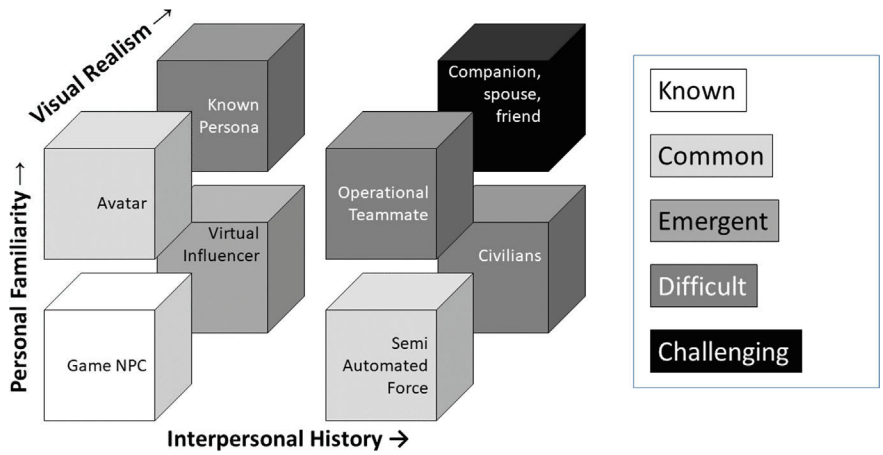

Fig. 1.1 Some virtual human dimensions and example applications

- **Visual Realism**: How important to the observer and the application are realistic human *physical* characteristics, such as shape, appearance, clothing, and faces? How important are accurate portrayals of body motions, postures, locomotion, gestures, object interactions, eye movements, and facial expressions?
- **Personal familiarity**: Does the observer *know* the virtual being? Is it a novel invention, where its background is unknown or constructed as needed? Is it a digital version of someone we've heard of or observed (perhaps through some visual media) in real life, such as an actor or sports figure? Is it someone we know quite well through first-hand experience?
- **Interpersonal history**: What *interactions* do we expect from this virtual human construction? Is the information flow one way from the virtual person to the observer, or do we share a common history, goal, or mission that requires communication, coordination, or intimacy?

In my diagram, the boxes represent selected generic applications for virtual humans. The shading indicates the approximate difficulty these applications present to the software developer. The absolute scale is mostly irrelevant; the relative position from "easy" to "hard" is important. As the applications move toward the more difficult extremes, the problems tend toward the frontiers of present software capabilities.

- A **Game NPC** is a "Non-Player-Character." These may be scripted individuals who respond in a predetermined fashion during use or game play. They may be carefully crafted and artist-designed to match the game genre. The game player may have no prior exposure to the NPC, so all information is controlled by the NPC's appearance, actions, inventory, and speech. A digital double of a motion picture actor for a film is

a high-quality, artistically created, non-interactive NPC. A digital museum guide, triggered interactively by observer location, might utter pre-recorded commentary. Recent developments in artificial intelligence natural language models are beginning to impart NPCs with relevant, real-time, unscripted conversational abilities.

- **Avatars** are computational embodiments of the user. At a minimum, avatar movements are controlled through some human initiative. Avatars may appear in human form or in any other realization suited to, and reflective of, the player's persona. Avatars typically mirror user desires through a suitable interactive control interface, such as a game controller, mouse and keyboard, or wearable sensors. Camera-based systems may detect and transform facial movements and gestures into the avatar's visual representation.
- A **semi-automated force** is a collection of non-player characters, usually in a conflict situation, who behave autonomously according to engagement rules. They may have explicit higher-level goals such as attack targets or threat avoidance. Once instructed, they work on their own. They may be allies or enemies. Their behavioral rules require knowledge of the environment and other agent activities. This information may be messaged through a digital communications medium or sensed by the agent's own synthetic perceptions.
- A **virtual influencer** is an enhanced avatar, usually with an emphasis on personal appearance, contemporary notions of attractiveness, and a cultural demographic. They are usually controlled by a real person through motion capture technology, so they can act human without requiring any automated behaviors. Body shape, facial expressions, and even hair movements are important visual characteristics. Real-time interaction is often unnecessary since they can be streamed online from pre-recorded video. Influencers are often invented personalities to avoid any overt prior knowledge of their history.
- A **known persona** is a virtual version of a real, well-known individual. They may be alive at present or a digital reconstruction of a historical figure. Their history, fame, achievements, and personality are familiar enough that departures from these norms destroy the purpose and effect of virtualization. Appearance, behavior, facial expressions, gestures, clothing, and speech must all conform to their known and actual ideals.
- An **operational teammate** is an individual virtual person who works with a user to achieve common goals. Their behavioral cues might arise from player directives, but the teammate also follows any engagement rules. Some of these rules may align with their companionship obligations on a game quest. There is a growing bond between the player and teammate in their shared history. Conversational agents are a compatible term for this category when speech and gestures are the primary communicative channels.
- **Civilians** are virtual people who populate a scene. They may be passive pedestrians or other animated (even nonhuman) scene occupants, such as vehicles, ostensibly piloted by people. They should obey traffic laws and navigate complex networks of paths

and roads. As virtual people, they can appear as members of a community, family, or cultural unit. The observer may feel a need to protect or avoid conflict with civilians because of these relationships. They should have enough visual realism to appear unique in dress or behavior but otherwise anonymous. The challenge lies in making these people appear to have roles, functions, and purpose within a larger environmental context.

- **Companions, spouses, and friends** are people we know very well in both public and private spheres. Virtual versions need to project mutual relationships based on intimacy, shared history, expectations, and interests. Interlocutors should expect empathy, emotional displays, voice and speech interaction, and a consistent personality. Private information must be recognized and respected during interactions. Appropriate and accurate appearance, movements, and shared interpersonal responses are crucial.

These characterizations can only address the principal attributes of these application spaces. Each has an extensive community, considerable published history, active users, observers, and collaborators. In this exposition, I cannot adequately survey the rapidly evolving state of these technologies, so I will adopt a different approach. The underlying commonality is the portrayal, in visual form, of human figures. Layered on top of that highly variable but fundamentally structured shape are its intrinsic physiological abilities, such as gestures, reaching and grasping, postural positioning, locomotion, eye movements, facial appearance, and speech. In addition, there are behaviors that organize abilities according to task achievement, communicative goals, environmental necessities, or personal maintenance. These factors, in turn, are moderated by individuality, emotion, personality, and interpersonal shared knowledge. Finally, there is the societal milieu that drives life goals, careers, ethics, action choices, and beliefs.

I have deliberately chosen to address the evolution of some of these digital human requirements as an origin story rather than a contemporary survey. I can provide considerable backstory, honor the real people who contributed to my own personal journey, and add the human experiences that often do not find their way into the technical publications that track progress. My contributions do not diminish those of many others. By the late 1980s, the virtual human animation field was already diverse, deep, and robust [2]. I merely know these stories well enough to add some life to virtual humans.

The graphic origin in Fig. 1.1 is the eponymous motivation of this narrative. Before I can address any of these dimensions or applications, I need to start with the even more basic question about why modeling humans came to be an integral part of my own life. Once aimed in that general direction, building basic human capabilities into a computer model occupied me, my colleagues, and my students for five decades. Unlike real humans, we can't start with a baby and watch it learn and grow. Moreover, unlike contemporary approaches to building computational analogs of language and image generation through machine learning and artificial intelligence methods, we did things incrementally by building models whose capabilities we understood. Our models could still exhibit interesting

and novel behaviors, surprise us (and others!) with their abilities and actions, and find meaningful and rewarding roles in the service of humanity. Building a digital human ab initio, was the precursor to designing more visual realism, building better communicative partners, and making them appear to share emotions and personalities. For most of this discussion, the terms *virtual* and *digital* are interchangeable. I am not a roboticist and do not build tangible humanoid entities, but our work has influenced others who create such inventions. I arrived at a succinct summary of my professional goals in the title of a talk I gave many times in the past three decades: "My lifelong quest to control virtual people."

Origin stories are often fascinating tales of connections, discoveries, and effort. In computer graphics, we celebrate the ideas and people who bring real and imaginary worlds into the realm of synthetic visuals. Tracing the origins of new ideas by infinite regression through the citation fields of technical papers is a classic exercise for graduate students embarking on a Ph.D. Most often, however, the processes by which big ideas develop and flourish are hidden in the personal experiences of the creators and builders along the evolutionary chain. This generative process is fraught with the same transformations and consequences contributed by natural biological evolution: an unexpected mating of ideas, a meaningful mutation of a core concept, a withering of less capable offspring, and even the butterfly effect of random environmental influences. As we look backward, we can better recognize this evolutionary process to seek or at least understand the triggering conditions. While in the midst of this evolution we may not clearly see those "aha" moments that will shape the future, they may be uncovered in retrospect to trace the crucial influences along the path.

I have been extremely fortunate to have experienced and contributed to the software development of digital humans. Whether they are avatars, conversational agents, game NPCs, or virtual influencers, they all must have origins in the universe of computer graphics. Indeed, the "DNA" of these human models encompasses a large array of computer graphics component technologies, including 3D modeling, animation techniques, motion capture, and interactive systems. Finding (or purporting to know) "the" path from these origins to the present would require writing a computer graphics history book. However, what I can do and will try to document here is trace an evolutionary path—and a personal journey—to elucidate the context and processes that led to the explosion of digital human software "DNA" which has enabled contemporary embodiments for virtual beings.

References

1. M. Mori. "The Uncanny Valley: The Original Essay." June 2012. https://spectrum.ieee.org/the-uncanny-valley.
2. R. Earnshaw, N. Magnenat-Thalmann, D. Terzopoulos and D. Thalmann. "Computer Animation for Virtual Humans." IEEE Computer Graphics and Applications. 1988, pp. 20–23.

Beginnings

2

In 1952, when I was four years old, my parents moved to California from their family roots in Philadelphia. I grew up in the San Francisco East Bay area, a baby boomer in a young community. Neither of my parents had any postsecondary education, and neither came from any academic or technical heritage. My father was a bomber copilot in southern Europe in World War II. After the war ended, he married my mother, and they moved to Los Angeles. Soon thereafter, I was born. After a failed attempt to start a photography business, they moved back to Philadelphia and had my sister Marilyn. When I was about four years old, we moved back to California again, this time settling in San Leandro in the Bay Area. My father mostly worked as a re-upholstery salesman. Soon, he switched to being a window salesman, taking advantage of the postwar construction boom. My mother was a part-time office secretary and homemaker. She did collect books (I never saw her reading them, however), and we had a decent, if small, library of novels, war documentaries, and art books. I think I eventually read all of them.

My father liked to build things. He was an avid balsa airplane modeler, at least until he had children. With the emergence of affordable electronic components, he ordered HeathKits to assemble his own stereo system. I got to help and was quite good at soldering, recognizing components, and interpreting resistor color code bands. My father built our own stereo speaker wood enclosures in his garage workshop. (I grew up believing that garages were never used for cars, only for power tools.) His airplane model prowess had morphed into full-scale cabinet-making. My exposure to the methodical, actively planned, error-critical, and stepwise execution of these creative projects were primordial seeds for my own algorithmic thinking. Rather than airplanes, my creative hobby was HO model railroading, something that exploited similar skills.

N. Badler, *On Raising a Digital Human*, Synthesis Lectures on Computer Science,
https://doi.org/10.1007/978-3-031-63945-6_2