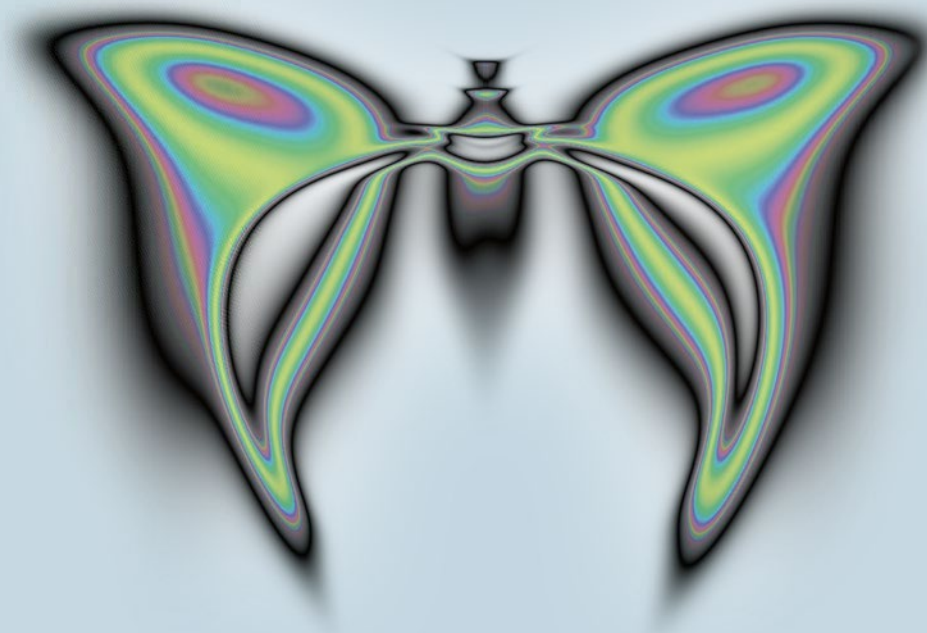


Kenneth O. Stanley · Joel Lehman

Why Greatness Cannot Be Planned

The Myth of the Objective



 Springer

Why Greatness Cannot Be Planned

Kenneth O. Stanley • Joel Lehman

Why Greatness Cannot Be Planned

The Myth of the Objective

 Springer

Kenneth O. Stanley
Department of EECS
Computer Science Division
University of Central Florida
Orlando, FL, USA

Joel Lehman
Department of Computer Sciences
The University of Texas at Austin
Austin, TX, USA

The image on the cover has been used with permission from the University of Central Florida Research Foundation, Inc. It was created using the www.picbreeder.org website of UCFRF.

ISBN 978-3-319-15523-4 ISBN 978-3-319-15524-1 (eBook)
DOI 10.1007/978-3-319-15524-1

Library of Congress Control Number: 2015932228

Springer Cham Heidelberg New York Dordrecht London
© Springer International Publishing Switzerland 2015

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

To Beth and Bennett
KS

To my parents
JL

Preface

This book was born from a radical idea about artificial intelligence (AI) that unexpectedly grew to be about much more. At first I was thinking only about AI algorithms, programs of concern mainly to computer scientists like me. Usually these algorithms have explicit goals and objectives that they're driven to achieve. But I began to realize that the algorithms could do amazing things even if they had no explicit objective—maybe even more amazing than the ones that did have an objective. Testing this idea led to some surprising experimental results, some of which are documented in this book. And that's pretty interesting—if you're a computer scientist.

But then something unusual happened. I started to realize that the insight wasn't just about algorithms, but also about life. And not just life, but culture, society, how we drive innovation, how we plan for achievement, our interpretation of biology—the list just kept expanding. If you don't understand why I say that's unusual, just consider how rare or bizarre it is for a computer algorithm to change how you think about life. After all, you don't fall into an existential crisis every time you boot up your laptop. The unexpectedly broad implications of the idea surprised me, so at first I kept them tucked quietly in the back of my head—but they kept getting louder.

As a professor, I'm sometimes invited to give public talks on my research. So as an experiment, when I spoke publicly about the idea I began touching on its relationship to our lives and society. And as I saw how people reacted and how much passion it provoked, the message took on a new life—because I could see it had meaning well beyond the field of its origin. So I realized someone had to write this book at least to try to communicate the novel insight. And that's why you're in for a unique experience in the pages ahead. There is a story here—a story about an idea in AI and how it grew into something bigger—but there's also a journey through a dizzying set of surprisingly broad implications for everything from personal dating, to the march of science, to the evolution of the human brain. I hope you'll enjoy this whirlwind tour across a landscape of once familiar concepts—but now observed through a psychedelic new lens.

One more important detail about the history of this project—it was really a two-person effort. From the earliest experiments to realizing the growing set

of implications, my coauthor Joel Lehman was instrumental all along the way. The ideas here are the result of years of exchange and debate between the two of us, and the book is therefore truly a joint effort. So from the first chapter onward, we'll be speaking with a single unified voice as we guide you through *Why Greatness Cannot Be Planned: The Myth of the Objective*.

Joel and I would both like to express gratitude to the institutions that supported this work: the University of Central Florida, the University of Texas at Austin, and the Santa Fe Institute, where I completed the book while on sabbatical. Every past and current member of the Evolutionary Complexity Research Group at UCF deserves special thanks for their input and ideas over the years. I am also indebted to Gary Leavens, the Chair of the Computer Science Division at UCF, who encouraged me to make this book a reality. This book also probably would not have happened without the motivation provided by early invitations to speak from Richard Gabriel of IBM Research (at the 2010 SPLASH conference) and Seung Chan Lim (Slim) of the Rhode Island School of Design in 2011. For creating the experiment that provided the initial inspiration for the ideas in this book, the entire Picbreeder team and later Picbreeder contributors deserve special mention: Jimmy Secretan (lead), Nick Beato, Adam Campbell, David D'Ambrosio, Adelein Rodriguez, and Jeremiah T. Folsom-Kovarik; Nazar Khan, Peter Matthews, and Jan Prokaj later contributed to researching Picbreeder's color extension.

* * *

The book is organized into two parts for the convenience of different readers. The first part of the book (the first nine chapters) constructs the main argument against objectives and provides general evidence for their cost in a number of areas of life and society. For readers interested in more elaborate implications of the myth of the objective in specific scientific fields (in particular biology and artificial intelligence), two additional case studies are included at the end of the book. That way, while you can absorb the main ideas from just the first nine chapters, these additional case studies provide further depth for those who desire more.

Santa Fe Institute, Santa Fe, NM
March 10, 2015

Kenneth Stanley

Contents

1	Questioning Objectives	1
2	Victory for the Aimless	13
3	The Art of Breeding Art	21
4	The False Compass	29
5	The Interesting and the Novel	39
6	Long Live the Treasure Hunter	55
7	Unshackling Education	65
8	Unchaining Innovation	79
9	Farewell to the Mirage	93
10	Case Study 1: Reinterpreting Natural Evolution	101
11	Case Study 2: Objectives and the Quest for AI	119
	Bibliography	137

Chapter 1

Questioning Objectives

Now, in his heart, Ahab had some glimpse of this, namely: all my means are sane, my motive and my object mad.

Herman Melville, *Moby Dick*

Imagine waking up without thinking about what you need to do today. Have you ever done that? Suppose then you go to work, but instead of holding the usual meetings and poring over benchmarks and milestones, your boss instead tells you to *do whatever you feel is most interesting*. What would you do? When you later check the news online, there is no talk of national testing standards or missed objectives for the economy. Strangely enough, teachers somehow still teach and money somehow still changes hands anyway. Maybe you post a profile on a dating website, but you leave blank all the answers on what you look for in a partner. Today you aren't looking for anything in particular, but that still doesn't stop you from looking. While you probably won't experience a day like this one anytime soon—a day without objectives—what would life be like if you could? Would it be confusing, difficult to navigate, too open-ended—or, quite the opposite, a lot *better*?

It's interesting that we rarely talk about the dominance of objectives in our culture even though they impact us from the very beginning of life. It starts when we're barely more than a toddler. That momentous first day that we enter kindergarten is the gateway to an endless cycle of assessment that will track us deep into adulthood. And all that assessment has a purpose—to measure our progress towards specific *objectives* set for us by society or by ourselves, such as mastering a subject and obtaining a job. The reality is that objectives lurk in the background from the earliest moments. But they only start there. Over the years they keep gathering steam, eventually enveloping practically everything we do.

If you want evidence for that, all you need to do is stop in at your local bookstore and take a look at the magazine rack. There you will be reminded that you might want to *change jobs, lose 15 pounds, start a company, find a date, get promoted, change your look, make a million dollars, buy a house, sell a house, or even complete a video game*. In fact, almost anything worth doing is expressed as an objective. Now, we're not suggesting by any means that all of these objectives are a waste of time. Of course, many are admirable. Others may be more frivolous, but whatever you may think of one objective or another, we rarely question the value of

framing all our pursuits with objectives. Can you imagine life with fewer objectives, or with no objectives at all? Would it lead to any good? Regardless of how you'd answer these questions, at the least they highlight how deeply our culture has come to revere objectives.

And it's not only about individual pursuits. While it's true that children are graded for their progress towards mastering a subject, the schools themselves are *also* graded. But in the case of the schools, their objective is to produce high student test scores. Even nations set objectives, like low crime, low unemployment, and low carbon emissions. A lot of effort and resources are spent measuring progress towards these objectives and others like them. There's an assumption behind these pursuits that isn't often stated but that few would think to question: We assume that any worthy social accomplishment is best achieved by *first* setting it as an objective and *then* pursuing it together with conviction. It makes you wonder, is there such a thing as accomplishment without objectives?

If you take a look at most professions, the answer would appear to be no. Take engineers for example. They typically set objectives through rigorous specifications. Then they continually measure how their prototypes compare to these specifications. Inventors are similar, conceiving an invention and then pursuing it as an objective. In the same spirit, scientists must come up with clear objectives to secure funding for their projects, which are then judged by how likely they are to achieve their objective. The list goes on. Investors set earnings objectives and corporations set profit objectives. Even artists and designers often mentally conceive a design and set its realization as their objective.

The weight of objectives on our thinking even impacts the way we talk about topics like animals in nature. After all, whenever evolution is discussed, we view animals through the lens of *survival and reproduction*, evolution's assumed objective. Even many *algorithms*, or programs that run inside computers, are designed to work towards some specific objective, such as finding the best search result or playing a better game of chess. In fact, these kinds of algorithms are quite common in artificial intelligence and machine learning. So it may come as no surprise that the term *objective function* is now practically a household name in such fields.

Maybe all this objective mania makes sense. At some level, we must believe it does to have allowed it to dominate our lives so completely. Or maybe it's something else—maybe we've become so used to objectives defining everything we do that we've forgotten that their value can even be questioned. Either way, there's a certain appeal to this kind of routine. The idea that all our pursuits can be distilled into neatly-defined objectives and then almost mechanically pursued offers a kind of comfort against the harsh unpredictability of life. There's something reassuring about the clockwork dependability of a world driven by tidy milestones laid out reliably from the starting line.

Though often unspoken, a common assumption is that the very act of setting an objective creates possibility. The very fact that *you put your mind to it* is what makes it possible. And once you create the possibility, it's only a matter of dedication and perseverance before you succeed. This can-do philosophy reflects how deeply

optimistic we are about objectives in our culture. All of us are taught that hard work and dedication pay off—if you have a clear objective.

Even so, perhaps you've felt qualms about this kind of thinking from time to time. It might sound good on the face of it, but what it *leads to* isn't always so comforting—legions of measurements, assessments, metrics—woven into every aspect of our lives. It's like we've become slaves to our objectives, toiling away towards impossible perfection. Objectives might sometimes provide meaning or direction, but they also limit our freedom and become straitjackets around our desire to *explore*. After all, when everything we do is measured against its contribution to achieving one objective or another, it robs us of the chance for playful discovery. So objectives do come with a cost. Considering that this cost is rarely discussed in any detail, maybe it's a good idea to look a little harder at what we're really giving up in exchange for such objective optimism.

But before we do, it's important that you know that we aren't pessimists. It may sound like this book is going in a cynical direction, but that isn't really true. In fact, we believe that human achievement has no limits. It's just that we're going to highlight a different path to achievement, without the need for objectives. There's a lot our culture has sacrificed in the name of objectives, and we're going to take it back. They've stolen our freedom to explore creatively and blocked us from serendipitous discovery. They ignore the value of following a path for its own uniqueness, rather than for where it may lead. The chapters ahead will show that great discoveries are lurking just beyond our fingertips, if only we can *let go* of the security blanket of the objective. Sometimes, the best way to change the world is to stop *trying* to change it—perhaps you've noticed that your best ideas are often those you were not seeking. We'll come back to this paradox later, but first let's think a little more about the way most people approach achievement today.

* * *

It usually starts with deciding on an objective. In a lot of professions, the first question you'll hear if you propose a new project is, "What is the objective?" People often say that a particular pursuit is not "well-defined" enough unless you can tie it to an objective. That's usually the only way to prove your idea is worth considering. For example, even if you have a strong hunch that combining two chemicals will lead to an interesting reaction, hardly any scientist will take you seriously unless you can *define* what that interesting reaction would be. Only then could you say that you have a clear objective, and a legitimate pursuit is underway.

Sometimes a word other than "objective" might be used, but usually it plays a similar role. For example, scientists often demand *hypotheses* from each other because they don't want to fund research only because it sounds "interesting." "What is the hypothesis?" they'll insist. It's similar to asking what is the objective. Without a hypothesis, an experiment is reduced to *mere speculation*, little more than child's play. The hypothesis, like the objective of any project you might want to pursue, certifies that it's worth pursuing. Even if it might not work out, it still provides a clear outcome that will allow people to judge your success or failure later.

This attitude doesn't only apply to business or science. It's personal too. If you join a dating site, you're supposed to know what you want and what you like, so that you can describe what you're "looking for." If you're starting college, you're supposed to figure out your major so that you have a clear body of knowledge to pursue. And as any high school student realizes who is applying to college, even hobbies are supposed to have *some* kind of purpose. "Staring at the wall" will not meet much respect on a college application, even if it's when you have your best ideas.

Whether you're an executive, a scientist, a student, or even just a single looking for a date, once your objective is defined usually the next step is to put all your energy into achieving it. In other words, we pull out all the stops in the pursuit of our goal. But there's actually one other ingredient in pursuing an objective that is not as obvious but almost always present: Usually progress towards the objective is somehow *measured*. This is where all the measurements and metrics of our culture come into play. The purpose of all the measuring is to help us figure out whether we're heading in the right direction. If things aren't going well, it gives us the chance then to switch directions. For example, in school the objective is to master a subject. Grades are a measure that tell us whether we're making progress in that direction. So if your grades in school are going down, you may need to change your approach to studying.

There's a useful word that comes from optimization theory that some scientists use to describe this idea of using a measure to help decide what to do next—they call the measure a *gradient*. It's basically a clue to which direction is right and which is wrong. We all quickly become very familiar with following such gradients, to the point that it almost becomes unconscious by the time we're adults. The "hot and cold" children's game is a perfect introduction to the idea: The young contender searches for a hidden treasure known only to the other players, who provide clues by saying "hotter" or "colder." The idea is so simple that children naturally follow the "gradient" of increasing temperature with little need for explanation (and definitely no need to learn anything about optimization theory). As long as the gradient is going up, chances are that success is coming closer. In a sense we never stop playing the hot and cold game. The process of setting an objective, attempting to achieve it, and measuring progress along the way has become the primary route to achievement in our culture.

Obviously we're going to raise a lot of questions about the benefits of objectives, but one important point is that we're mainly focused on *ambitious* objectives—those whose achievement is anything but certain. One of the reasons that objectives aren't often questioned is that they work perfectly well for more modest pursuits. If a manufacturer decided to increase efficiency by 5%, no one would be shocked if it succeeded. A software company upgrading its product from version 2.0 to version 3.0 is similarly likely to succeed, as happens all the time. Everyday successes like these mislead us into believing that setting objectives works well for almost *everything*. But as objectives become more ambitious, reaching them becomes less promising—and that's where the argument becomes most interesting.

Some objectives are anything but certain. Medical researchers have yet to cure cancer. And it's hardly clear that computer scientists will succeed in creating a convincingly-human artificial intelligence any time soon. It would be nice to have a limitless source of energy with no risk or environmental impact whatsoever, but who knows when that will happen. Even fully immersive holographic television, which you must admit would be fun though perhaps less noble, is nowhere in sight. Maybe there is even a beautiful new genre of music waiting to be found, so enchanting that all of humanity would succumb to its spell, if only the right artist can find it. Or, if you want to get *really* ambitious, how about time travel or teleportation? You might even have some big ones of your own, like making a billion dollars. Of course some ambitious objectives might turn out impossible. But *some* are possible, and if we could even come close to achieving them, the world would surely change for the better.

The question then is, *how will we ever get there?* How can we achieve not just improving our golf swing, but our *dreams*? These are the kinds of questions where objectives offer less assurance. After all, even if time travel is theoretically possible, the best use of national resources at this moment is surely not to reroute trillions of dollars into an all-out effort to build a time machine. But why not? Is setting the objective not the first step towards success? Or would pursuing time travel be our Moby Dick, distracting us from what really matters in life?

* * *

Why is it so hard to achieve ambitious objectives? To get a handle on the problem, think about all the things that are *possible*. As an example, consider all possible *images*. Everything you've ever seen falls into this class, but so do many things you've never seen and will never see. Now consider this: A minuscule proportion of all possible images are great masterpieces like Leonardo da Vinci's *Mona Lisa* or Vincent Van Gogh's *The Starry Night*. These are objectives that are hard to achieve. A larger proportion of possible images are recognizable but less inspiring than the masterpieces, like pictures of acquaintances and everyday objects. Of course, the vast majority of possible images are of no interest whatsoever—like television screens tuned to the wrong channel, just random static. One exciting aspect of this vast set of all possible images is that some of the great masterpieces *do not yet exist* because no one has painted or drawn them yet. Put another way, these future masterpieces are not yet *discovered*.

It's useful to think of achievement as a process of discovery. We can think of painting a masterpiece as essentially discovering it within the set of all possible images. It's as if we are *searching* through all the possibilities for the one we want, which we call our objective. Of course, we're not talking about search in the same casual sense in which you might search for a missing sock in the laundry machine. This type of search is more elevated, the kind an artist performs when exploring her creative whims. But the point is that the familiar concept of search can actually make sense of more lofty pursuits like art, science, or technology. All of these pursuits can be viewed as searches for something of value. It could be new art, theories, or inventions. Or, at a more personal level, it might be the search for the right career.