Chapter 4
Decoding What is Good in Code:
Toward a Metaphysical Ethics of Unicode

Jennifer Helene Maher
University of Maryland, Baltimore County, USA

ABSTRACT

Programming benefits from universal standards that facilitate effective global transmission of information. The Unicode Standard, for example, is a character encoding system that aims to assign a unique number set to each letter, mark, and symbol in the world's various written systems, including Arabic, Korean, Cherokee, and even Cuneiform. As the quantity of these numerical encodings grow, the differences among the written systems of natural languages pose increasingly little consequence to the artificial languages of both programmers and machines. However, the instrumental, technical effects of Unicode must not be mistaken as its only effects. Recognized as a metaphysical object in its own right, Unicode, specifically, and code, generally, creates a protocol for the actualization of moral and political values. This chapter examines how Microsoft's inclusion and then deletion of the Unicode encodings U+5350 and U+534D in its Office Bookshelf Symbol 7 font illustrates how technically successful coding can be rhetorically buggy, meaning it invokes competing ethical values that, in this case, involve free speech, anti-Semitism, and Western privilege.

INTRODUCTION

One of the greatest challenges to global literacy in the digital age has proven to be the diversity that exists among natural languages. Of the approximately 7,000 known living natural languages, approximately 80 also exist as a written system. Consequently, one of the most elemental challenges to the digital transmission of information is the character variation among these systems. To process information digitally necessitates a character encoding systems that allows for the unambiguous translation of not only alphabetic (e.g., Latin), logographic (e.g., Chinese) and syllabic (e.g., Cherokee) characters but also characters necessary for punctuation, ideograms and control at both the level of the human-readable artificial programming languages and the machine-readable binary code of 1's and 0's. But creating a standard character encoding system for even one written language is challenging. IBM computer scientist Robert Bemer described the numerical encoding of the relatively simple, Latin alphabet-based Standard American English in the period before the 1960 development of the character encoding system American Standard Code for Information Interchange (ASCII) as nothing short of "the Babel of internal computer codes."

Certainly the development of ASCII had done much to improve character encoding functionality in the Latin-based alphabetic system. However, as information exchange went global via the Internet, the challenge became not simply how to how to encode characters and symbols within particular alphabetic systems but how to encode across those systems. As Joseph D. Becker (1988), of Xerox Corporation, explained in his seminal paper "Unicode 88":

[T]he people of the world need to be able to communicate and compute in their own native languages, not just in English. Text processing systems designed for the 1990s and the 21ˢᵗ century must accommodate Latin-based alphabets

for European language such as French, German, and Spanish; and also major non-Latin alphabets such as Arabic, Greek, Hebrew, and Russian; and also "exotic" scripts of growing importance such as Hindi and Thai; not to mention the thousand of ideographic characters used in writing Chinese, Japanese, and Korean. (p. 1)

To address this need, the non-profit Unicode Consortium was formed in 1991 and described its mission as arising out of a need to develop a set of universal standards and specifications so as to enable "people around the world to use computers in any language." Published in the same year, the *Unicode Standard*, according to the Consortium's website, offered "a unique number for every character, no matter what the platform, no matter what the program, no matter what the language." Although often invisible to non-technical end-users unaware of this standard, the Consortium makes clear the importance of the *Unicode Standard* to the most quotidian practices of the Internet: "Our freely-available specifications and data form the foundation for software internationalization in all major operating systems, search engines, applications, and the World Wide Web."

Unicode is undoubtedly an important element to the successful workings of Internet communication. However, at present, Unicode is too often considered only in terms of a technical standard that maximizes efficiency and effectiveness. The numeric translation of characters and symbols appears to facilitate the boiling down of the means of written communication to its most powerful essence as series of 1's and 0's read by the computing machine, a reduction that frees language from the complex, social politics that circulate through it. To think of language then as essentially just a collection of discreet units of information belies the powerful dynamics that are masked rather than resolved through Unicode. In this chapter, I offer an example of what it means to think of the object of code, generally, and Unicode, specifically as still part of an ancient, human tradition concerned with

Related Content

Social Media Marketing In Emerging Economies: Case Study of Three Indian Firms
Madhupa Bakshiand Prashant Mishra (2014). *Digital Arts and Entertainment: Concepts, Methodologies, Tools, and Applications (pp. 627-649).*
www.irma-international.org/chapter/social-media-marketing-in-emerging-economies/115038

An Integrated Approach to Narrative Generation: From Mishima and Kabuki to Narrative Generation Systems
Takashi Ogata (2018). *Content Generation Through Narrative Communication and Simulation (pp. 49-147).*
www.irma-international.org/chapter/an-integrated-approach-to-narrative-generation/200229

An Evaluation of the Usability of a DMO's Digital Marketing Strategy: The Case of the Center of Portugal
Bernardo Borgesand Rui Costa (2021). *Impact of New Media in Tourism (pp. 88-101).*
www.irma-international.org/chapter/an-evaluation-of-the-usability-of-a-dmos-digital-marketing-strategy/272420

Young Adult Use of Sexting
Rachel E. Riggsand Narissra M. Punyanunt-Carter (2020). *Young Adult Sexuality in the Digital Age (pp. 1-12).*
www.irma-international.org/chapter/young-adult-use-of-sexting/255497

Exploring Dimensions of the Media Dream: Functional Context in Collective Personae
Rollin McCratyand Stephen Brock Schafer (2016). *Exploring the Collective Unconscious in the Age of Digital Media (pp. 1-39).*
www.irma-international.org/chapter/exploring-dimensions-of-the-media-dream/145258