Anyone who has ever prepared a critical edition of any text from more than one or two manuscripts, whether in Armenian or any other language, will be familiar with the vast amount of laborious, repetitive, and careful work it requires. The texts must be transcribed, taking care not to introduce the sorts of copying errors that make critical editions necessary to begin with. A “base” text must be chosen, and subsequent texts compared with it; the granularity of the comparison, and the consequent quality of the eventual edition, is the direct result of a scrupulous attention to detail at this stage that taxes the abilities of almost all editors. Next, the transcribed and collated texts must be analyzed for their relationships to each other and subjected to critical analysis. Here the editor must again display prodigious attention to detail and find a way to ensure consistency in his or her editing decisions throughout a process that may stretch over months or years. Somewhere along the way, the discovery of a new manuscript or the sudden realization that one text is rather more or less important than previously assumed will inevitably compel the editor to review, revise, or completely re-do some portion of this painstaking work. When this happens, the editor must accept his or her fate, taking care to ensure that no new errors creep in through frustration or impatience. Little wonder that so many scholars, having sent away the final proofs of a critical edition for publication, promise themselves “Never again!”

It need not be like this. If there is any advantage to living in the age of computers, it is that there should be ever less need for repetitive and codifiable tasks to be undertaken by human scholars. In this chapter we will examine computational methods that are helpful to the philologist, indicating what is possible today as well as promising directions for the future.

Greetham (1992) describes two “apparently mutually exclusive” schools of text-criticism methodology: that which attempts to reproduce the presumed intention of the original author, and takes a belletristic approach to the task, and that which regards editing as a task partially dependent on verifiable scientific principles. The debate has usually been framed as a competition between editorial taste and scientific method as the preferred way to arrive at a “best” text from the available evidence. Digital philology, with the prominence of computer programs and mathematical methods, would seem to belong to the latter camp, and has indeed been treated with suspicion by scholars who reject the notion of a machine-produced edition (Robinson 2004). In fact, there is
no such thing as a machine-produced critical edition, and this is not the goal of digital philology. Any method of text edition, no matter the school, has an element of the meticulous and repetitive about it. The aim of digital methods is to assume the burden of these repetitive and predictable tasks, leaving the human editor free to apply his or her interpretation and judgment to the substantial, and more interesting, questions that remain.

By far the primary advantage of the digital age is that rigorous, codifiable, and repetitive work may be delegated to the computer, which after all does not become bored or careless in its work. There is a great deal of such repetitive work in the process of text critical edition. To take full advantage of the computer’s capacity for handling the codifiable tasks, the steps of creating a digital edition must be separated into those that are necessarily manual and those that can be automated. An emerging “digital” method of text criticism is best broken into several distinct steps, as suggested by Robinson (2004):

- Transcription
- Collation
- Analysis (e.g. stemmatic analysis)
- Edition
- Publication

All of these must be done for any text edition; traditional methods, however, usually combine transcription and collation, and stemmatic analysis has often been attempted before transcription even begins. In our sequence of steps in digital philology above, transcription refers to the process of replicating the content of a manuscript text into a computer file for later manipulation and display. This is possibly the most important step in the process, as an accurate set of transcriptions is the basis for a good critical edition. Collation is the process of reconciling these transcribed texts, finding the correspondences and variations between them. Analysis refers to the use of computer programs to help the editor draw conclusions about the transcribed and collated text. Several forms of analysis may be appropriate at this stage, such as linguistic analysis or stylistic analysis for author attribution; here we will focus particularly on stemmatic analysis. This is the creation of a hypothetical “family tree” to describe the relationships between the extant manuscript texts. Edition is the step during which the collated texts are reviewed for their correspondences and differences, and the editor applies his or her scholarly judgment to the question of what, from the various alternatives (including, on occasion, emendation), the best text is. Publication is the final product – a printed book