Using artificial intelligence to fill in gaps in ancient texts

July 29 2019, by Nicola Holzapfel

Credit: Ludwig Maximilian University of Munich

Enrique Jiménez makes use of artificial intelligence to fill in the many tantalizing gaps in the surviving texts written on clay tablets by scribes in early urban societies 3000 years ago.

A single character can often make the difference between a clear-cut reading and a puzzling conundrum. Enrique Jiménez, a specialist in

Ancient Near Eastern Studies, is engaged on the [reconstruction](#) of the beginnings of world literature. Mesopotamia was home to the first true civilization, and a rich and bilingual literary culture had appeared there by around the year 2500 BCE. Its texts were written in cuneiform script on [clay tablets](#) in Sumerian and Akkadian. These texts survive only in fragmentary form, and many fragments measure only a few centimeters. The challenge for Jiménez and his colleagues is to identify pieces that belong together –the script itself is polyvalent and the semantics are highly context dependent. "Having to abandon work on a text because a single sign is missing is incredibly frustrating, especially when you know that there has to be a fragment somewhere that fills the gap."

Enrique Jiménez has set himself the goal of filling these agonizing gaps with the help of artificial intelligence, and he received a Sofia Kovalevskaya Prize from the Alexander von Humboldt Foundation to help him to achieve this aim. The award is intended to enable junior academics to undertake innovative research projects in a field of their own choice. Jiménez chose to pursue his project at LMU, where he holds a dual appointment in the Institute of Assyriology and at the Department of Ancient History, where his research group on Electronic Babylonian Literature is based. He and his team have begun to assemble a "Fragmentarium," a database in which all the text fragments in world museums will find their place. The project began in May 2018, and the database now contains some 10,000 fragments. "I hope to get up to about 15,000 by the end of the year," he says. This will allow all previously unpublished fragments from the library of the Assyrian King Assurbanipal to be located in less than a second.

40,000 fragments in the British Museum

Assurbanipal built his library in his capital city of Nineveh 2700 years ago, with the intention of bringing together the collected knowledge and wisdom of his time. After his death, the library was destroyed by the

Babylonians who conquered the city in the year 612 BCE and its holdings were shattered. Most of the fragments recovered by archaeologists are now in the British Museum in London. About half of these have yet to be deciphered—and Enrique Jiménez has taken on this task as part of his project, in cooperation with the Museum. In May, two members of the Museum's staff began to photograph the remaining collections of clay tablets held by the institution, and Jiménez and his group will have 40,000 photographs of fragmentary cuneiform tablets to work on over the next 5 years.

For years, Jiménez has been exploring the opportunities opened up by digital media. To begin with, he wrote simple programs himself. Now he works closely with data scientists. During his time at Yale University, he was involved in the Cuneiform Commentaries Project, whose was to publish online commentaries on Ancient Mesopotamian literature. He is also making the results of his research available on a dedicated website. He is currently engaged in training the program to 'read' Akkadian and recognize all the possible readings of the signs used in Akkadian texts. "The goal is to enable the program to recognize sequences of characters, which it can automatically slot into the appropriate context. This is much easier for a computer than it is for us humans. A computer can take all imaginable possibilities into account at once, and automatically compare new fragments with previously stored texts." The immediate goal is to complete or reconstruct texts that are already known in part, and there are plans to publish translations into English, German and Arabic.

Jiménez himself focuses on literary texts that were in circulation in what is now Iraq in the first millennium BCE—when Akkadian literature was in a flourishing state—and on texts that are written in poetic forms. Over the course of the project, Jiménez intends to work on a text corpus comprising approximately 10,000 lines. In this endeavor, he is collaborating with Dr. Anmar A. Fadhil of Baghdad University, who curates an archive of literary tablets found in the famed library of Sippar

in northern Babylonia. The Greek historian Berossos, whose life spanned the turn of the 4th to the 3rd century BCE, explains how it allegedly came into being: When Noah learned that the Deluge was imminent, and before he boarded the Ark, he was said to have buried the whole corpus of Mesopotamian writings in the city of Sippar before boarding the Ark. Sippar was located near the site of present-day Baghdad, but its library was only discovered in 1986. Among the texts recovered were the epic "Enuma eliš" (the Babylonian version of the Creation, and "a very beautiful text," says Jiménez) and an account of the Great Flood. "On the one hand, these texts are very different from the stories that are familiar to us. On the other, there are a great many parallels, including thematic parallels, between the different versions. Together with our Iraqi colleagues, we plan to publish the most important literary tablets. One of these is a fragment of the poem on the "Sufferings of the Righteous Man," the story of the Babylonian Job."

The poem tells the story of a just man, who nevertheless undergoes repeated punishments. At the midpoint of the text, the protagonist is first punished, then pardoned, and receives news of the pardon in a dream. Jiménez and Fadhil were able to assign a short text fragment from the library of Sippar to the dream episode in the poem, thus closing a gap in the previously known version. "We are now able to decode this passage, which nobody has been able to read for the past 2000 years. We have reconstructed a literary text that was recognized as a classic in antiquity, and rescued it for present-day readers."

Jiménez also identified fragments of a hymn addressed to the Babylonian god Marduk. "The texts that I work on are poetic texts. Akkadian poems are written in lines that have a clearly marked caesura in the middle. It is very difficult to translate this poetic structure into German." In scholarly publications, the texts are normally translated word-for-word into English. Enrique Jiménez has translated the reassembled and now intact lines into German, and the English version

given below is a free rendering of that [text](#) (the newly discovered phrases are shown in italics):

"Marduk, your anger is like a raging flood,
But in the morning, your mercy gives consolation to *your victim*,
The fearsome and terrifying storm *has abated*,
Where the waves thundered, *the shore is now calm*,
The dark threatening cloud *you have now illuminated*.
Where the searing scorching wind blew, *your bright day now reigns*."

Researchers can learn a lot from these texts about everyday life in Ancient Mesopotamia—about the objects they handled each day and about social life in general. "The Sufferings of the Righteous Man" describes how the protagonist awakes and speaks with his entourage—and the passage shows that he has a servant whose duties included waking his master up in the morning.

"In terms of its function too, Babylonian literature is not that far removed from our own," says Jiménez. The classical works were constantly copied, largely because they served as sources of exemplary passages for the teaching of reading and writing in schools. But the Creation Epic was recited in public at the festival traditionally held to celebrate the coming of the New Year. The Babylonian classics also suffered a fate that remains familiar even today: they were often parodied. "These parodies tell us a lot about the relationship between the people of Mesopotamia and their own traditions. For example, one line in the "Sufferings of the Just Man" was modified slightly so as to turn it into a joke. We have a very similar attitude to our literary traditions in the modern world. On the one hand, the classics are highly respected, and at the same time one can play around with them."

Unfortunately, there are many gaps in the parodies that have been studied, because so many of the tablets have suffered damage at one

time or another. "I am confident that, with the help of the computer, we will be able to reconstruct the whole of Akkadian literature in the coming 40 to 50 years. For us, digital media are simply tools: We are not working on algorithms in the abstract. We want to use them to establish the recorded texts. Without knowledge of its literary classics, it is impossible to understand any civilization."

Provided by Ludwig Maximilian University of Munich