Misinterpretations of Kolmogorov complexity

Kalevi Kolttonen
kalevi@kolttonen.fi

2018-10-18
v1.01

1 Introduction

In my earlier article *Metalevels in thinking* [1] I argued that not being careful with always maintaining e.g. a clean separation of object language and metalanguage can lead to significant confusion. That confusion can, in turn, cause us to draw conclusions that *do not* logically follow from our premises.

I do not bother to argue here for the crucial importance of avoiding incorrect inference. Instead I will take it as granted that the readers of this paper are already aware of this basic principle behind valid reasoning.

2 Background information

In *Metalevels in thinking* I mentioned Quine's example of a word ('Boston') and the object it refers to (city of Boston). I also warned about misinterpreting *concatenation* function to apply to numbers, proving with the use of hexadecimal notation (i.e. base16) that *concatenation* can never apply to numbers themselves, but only to the *names* of numbers.

Most philosophers know that *Plato* thought about the nature of numbers, and about all abstract objects in general, in Greece already over 2400 years ago. So that topic is far from new. Although I consider Plato's *The Republic* to describe a pretty horrible "ideal state", it must be admitted that Plato *never* confused numbers as objects with the names of numbers.

That is why he postulated a special universe for abstract objects, but it is of course probably true that that assuming the existence of such a universe quite possibly leads to infinite regression. In any case it is unfortunate that Plato's basic ideas are not always known and properly understood by mathematicians. Sometimes even philosophers have problems grasping them.

3 Dangers of not identifying abstraction levels

I do not want to mention any particular researchers here, but if you do some investigations on your own, I am sure you will soon admit that some people studying Kolmogorov complexity [2] *are* misinterpreting the implications of their studies.

As of now (October 18th, 2018) indeed even the Wikipedia article on Kolmogorov complexity makes a false claim:

The Kolmogorov complexity can be defined for any mathematical object, but for simplicity the scope of this article is restricted to strings.

Too many mathematicians first occupy themselves with reasoning about names that refer to, for example, natural numbers. After their possibly quite correct reasoning, they make the confused leap to *a wrong level of abstraction.* For example, they could mistakenly claim that they have proven something about *the nature of natural numbers.*

This is an embarrassment, because in science, and even in our everyday life, our reasoning must never be invalid. In other words, our conclusions must always logically follow from our premises in whatever logical system we happen to use.

4 Kolmogorov complexity never applies to numbers

Using hexadecimal notation again like we did in *Metalevels of thinking*, it is easy to show that studying e.g. a string like:

$aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa$

can lead us to say that this string can be expressed in a more compact way.

We see that it has 60 $'a'$ characters, so we could formulate a compression system using a shorter string like:

60a

In this example that string would mean that character $'a'$ should be repeatedly concatenated to the empty string 60 times. So far so good.

Sadly the confused researchers would now claim that this proves something about the *actual number object* referred to by this string.

Let's create a trivial Python 3 program to convert base16 to the more familiar base10:

```
#!/usr/bin/python3 -tt

print('%d' % 0xaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa)
```

Running this program produces the following output using base10:

1177898043185589553055531667161945677218322597917079305414404134195079850

By now we should realize that our earlier compressed string "60a" is no longer applicable after conversion from base16 to base10. The name of our number looks significantly different from the 60 consequtive 'a' characters that we had in base16.

But it is rather evident that both names must refer to the *same number*, i.e. the same object, perhaps located in Plato's universe of abstract objects.

If this were not so, we would have to assume the existence of different number objects *for each base* we could have. Obviously we can express the names of numbers in *any base* we wish, as long as we invent enough new symbols to satisfy our new base's needs. In other words, base10000000 would require one million different symbols, and so on.

Clearly this violates the well-known philosophical principle called Occam's razor. According to that principle, we should always choose an ontology that is *as simple as possible.* If we were to interpret each base as having their own, unique number objects, then this generally accepted scientific principle would be grossly violated.

In addition, it would be totally mysterious why the familiar mathematical functions such as addition always yield the same results after the results are converted into a uniform base.

References

[1] Metalevels in thinking. https://kolttonen.fi/philosophy/metalevels_in_thinking/met
Accessed: 2018-10-18.

[2] Wikipedia kolmogorov complexity. https://en.wikipedia.org/wiki/Kolmogorov_complex
Accessed: 2018-10-18.