



I'm not a robot



Continue

Python list comprehension loop order

When working with lists in Python, you'll likely often find yourself in situations where you'll need to translate values from one list to another based on specific criteria. Generally, if you're working with small datasets, then using for loops instead of list comprehensions isn't the end of the world. However, as the sizes of your datasets start to increase, you'll notice that working through lists one item at a time can take a long time. Let's generate a list of ten thousand random numbers, ranging in value from one to a million, and store this as num_list. We can then use a for loop and a list comprehension to generate a new list containing the num_list values greater than half a million. Finally, using %timeit, we can compare the speed of the two approaches: List comprehensions are useful and can help you write elegant code that's easy to read and debug, but they're not the right choice for all circumstances. They might make your code run more slowly or use more memory. If your code is less performant or harder to understand, then it's probably better to choose an alternative. Comprehensions can be nested to create combinations of lists, dictionaries, and sets within a collection. For example, say a climate laboratory is tracking the high temperature in five different cities for the first week of June. The perfect data structure for storing this data could be a Python list comprehension nested within a dictionary comprehension: >>>> cities = ['Austin', 'Tacoma', 'Sacramento', 'Charlotte'] >>> temps = {city: [0 for _ in range(7)] for city in cities} >>> temps = {city: [0 for _ in range(7)] for city in cities} Nested lists are a common way to create matrices, which are often used for mathematical purposes. Take a look at the code block below:

```
>>>> matrix = [[i for i in range(5)] for _ in range(6)] >>> matrix [[0, 1, 2, 3, 4], [0, 1, 2, 3, 4], [0, 1, 2, 3, 4], [0, 1, 2, 3, 4], [0, 1, 2, 3, 4]] The outer list comprehension [...] for _ in range(6)] creates six rows, while the inner list comprehension [i for i in range(5)] fills each of these rows with values. So far, the purpose of each nested comprehension is pretty intuitive. However, there are other situations, such as flattening nested lists, where the logic arguably makes your code more confusing. Take this example, which uses a nested list comprehension to flatten a matrix: >>> matrix = [..., [0, 0, 0, ... [1, 1, 1], ... [2, 2, 2], ... ] >>> flat = [num for row in matrix for num in row] >>> flat [0, 0, 0, 1, 1, 2, 2, 2] The code to flatten the matrix is concise, but it may not be so intuitive to understand how it works. On the other hand, if you were to use for loops to flatten the same matrix, then your code will be much more straightforward: >>>> matrix = [..., [0, 0, 0, ... [1, 1, 1], ... [2, 2, 2], ... ] >>> flat = [] for row in matrix: for num in row: ... flat.append(num) ... >>> flat [0, 0, 1, 1, 2, 2, 2] Now you can see that the code traverses one row of the matrix at a time, pulling out all the elements in that row before moving on to the next one. While the single-line nested list comprehension might seem more Pythonic, what's most important is to write code that your team can easily understand and modify. When you choose your approach, you'll have to make a judgment call based on whether you think the comprehension helps or hurts readability. A list comprehension in Python works by loading the entire output list into memory. For small or even medium-sized lists, this is generally fine. If you want to sum the squares of the first one-thousand integers, then a list comprehension will solve this problem admirably: >>>>> sum([i * i for i in range(1000)]) 332833500 But what if you wanted to sum the squares of the first billion integers? If you try to do this on your machine, then you may notice that your computer becomes non-responsive. That's because Python is trying to create a list with one billion integers, which consumes more memory than your computer would like. Your computer may not have the resources it needs to generate an enormous list and store it in memory. If you try to do it anyway, then your machine could slow down or even crash. When the size of a list becomes problematic, it's often helpful to use a generator instead of a list comprehension in Python. A generator doesn't create a single, large data structure in memory, but instead returns an iterable. Your code can ask for the next value from the iterable as many times as necessary or until you've reached the end of your sequence, while only storing a single value at a time. If you were to sum the first billion squares with a generator, then your program will likely run for a while, but it shouldn't cause your computer to freeze. The example below uses a generator: >>>>> sum(i * i for i in range(1000000000)) 3333333283333333500000000 You can tell this is a generator because the expression isn't surrounded by brackets or curly braces. Optionally, generators can be surrounded by parentheses. The example above still requires a lot of work, but it performs the operations lazily. Because of lazy evaluation, values are only calculated when they're explicitly requested. After the generator yields a value (for example, 567 * 567), it can add that value to the running sum, then discard that value and generate the next value (568 * 568). When the sum function requests the next value, the cycle starts over. This process keeps the memory footprint small. map() also operates lazily, meaning memory won't be an issue if you choose to use it in this case: >>>>> sum(map(lambda i: i * i, range(1000000000))) 3333333283333333500000000 It's up to you whether you prefer the generator expression or map(). So, which approach is faster? Should you use list comprehensions or one of their alternatives? Rather than adhere to a single rule that's true in all cases, it's more useful to ask yourself whether or not performance matters in your specific circumstance. If not, then it's usually best to choose whatever approach leads to the cleanest code! If you're in a scenario where performance is important, then it's typically best to profile different approaches and listen to the data. timeit is a useful library for timing how long it takes chunks of code to run. You can use timeit to compare the runtime of map(), for loops, and list comprehensions: >>>>> import random >>> import timeit >>> TAX_RATE = .08 >>> txns = [random.randrange(100) for _ in range(1000000)] >>> def get_price(txn): ... return txn * (1 + TAX RATE) ... >>> def get_prices(txns): ... return list(map(get_price, txns)) ... >>> def get_prices_with_comprehension(): ... return [get_price(txn) for txn in txns] ... >>> def get_prices_with_map(): ... prices = [] ... for txn in txns: ... prices.append(get_price(txn)) ... return prices ... >>> timeit.timeit(get_prices_with_map, number=100) 2.055437097998566 >>> timeit.timeit(get_prices_with_comprehension, number=100) 2.3982384680002724 >>> timeit.timeit(get_prices_with_loop(), number=100) 3.0531821520007725 Here, you define three methods that each use a different approach for creating a list. Then, you tell timeit to run each of those functions 100 times each. timeit returns the total time it took to run those 100 executions. As the code demonstrates, the biggest difference is between the loop-based approach and map(), with the loop taking 50% longer to execute. Whether or not this matters depends on the needs of your application. Suppose, we want to separate the letters of the word human and add the letters as items of a list. The first thing that comes in mind would be for loop. Example 1: Iterating through a string Using for Loop h_letters = [] for letter in 'human': h_letters.append(letter) print(h_letters) When we run the program, the output will be: ['h', 'u', 'm', 'a', 'n'] However, Python has an easier way to solve this issue using List Comprehension. List comprehension is an elegant way to define and create lists based on existing lists. Let's see how the above example, a new list is assigned to variable h_letters, and list contains the items of the iterable string 'human'. We can print() function to receive the output. Syntax of List Comprehension [expression for item in list] We can now identify where list comprehensions are used. If you noticed, human is a string, not a list. This is the power of list comprehension. It can identify when it receives a string or a tuple and work on it like a list. You can do that using loops. However, not every loop can be rewritten as list comprehension. But as you learn and get comfortable with list comprehensions, you will find yourself replacing more and more loops with this elegant syntax. List Comprehensions vs Lambda functions List comprehensions aren't the only way to work on lists. Various built-in functions and lambda functions can create and modify lists in less lines of code. Example 3: Using Lambda functions inside List comprehensions = list(map(lambda x: x, 'human')) print(letters) When we run the program, the output will be: ['h', 'u', 'm', 'a', 'n'] However, list comprehensions are usually more human readable than lambda functions. It is easier to understand what the programmer was trying to accomplish when list comprehensions are used. Conditionals in List Comprehension List comprehensions can utilize conditional statements to modify existing list (or other tuples). We will create list that uses mathematical operators, integers, and range(). Example 4: Using if with List Comprehension number_list = [x for x in range(20) if x % 2 == 0] print(number_list) When we run the above program, the output will be: [0, 2, 4, 6, 8, 10, 12, 14, 16, 18] The list, number_list, will be populated by the items in range from 0-19 if the item's value is divisible by 2. Example 5: Nested IF with List Comprehension num_list = [y for y in range(100) if y % 2 == 0 if y % 5 == 0] print(num_list) When we run the above program, the output will be: [0, 10, 20, 30, 40, 50, 60, 70, 80, 90] Here, list comprehension checks: Is y divisible by 2 or not? Is y divisible by 5 or not? If y satisfies both conditions, y is appended to num_list. Example 6: if...else With List Comprehension obj = "Even" if y % 2 == 0 else "Odd" for i in range(10): print(obj) When we run the above program, the output will be: ['Even', 'Odd', 'Even', 'Odd', 'Even', 'Odd', 'Even', 'Odd'] Here, list comprehension will check all 10 numbers from 0 to 9. If it is divisible by 2, then Even is appended to the obj list. If not, Odd is appended. Nested Loops in List Comprehension Suppose, we need to compute the transpose of a matrix that requires nested for loops. Let's see how we do the same thing normally for loop first. Example 7: Transpose of Matrix using Nested Loops transpose = [[0, 1, 2, 3, 4], [4, 5, 6, 7, 8]] for i in range(5): for j in range(5): print(transpose[i][j]) When we run the above program, the output will be: [[1, 2, 3, 4], [4, 5, 6, 7, 8]] In above program, we have a variable matrix, which have 4 rows and 2 columns. We need to find transpose of the matrix. For that, we used list comprehension. **Note: The nested loops in list comprehension don't work like normal nested loops. In the above program, for i in range(2) is executed before row[i] for row in matrix. Hence, at first, a value is assigned to i then item directed by row[i] is appended in the transpose variable. Key Points to Remember List comprehension is an elegant way to define and create lists based on existing lists. List comprehension is generally more compact and faster than normal functions and loops for creating list. However, we should avoid writing very long list comprehensions in one line to ensure that code is user-friendly. Remember, every list comprehension can be rewritten in for loop, but every for loop can't be rewritten in the form of list comprehension.
```


