Examining the Identity of a Sliced Python Object

David Prager Branner
Hacker School and PyGotham, New York
20140814 and 20140816

Slicing is a quick way to make a deep copy

Slicing is a quick way to make a deep copy — a copy of the actual values

Slicing is a quick way to make a deep copy — a copy of the actual values — of a sequence such as a list

$$>>> a = [1, 2, 3]$$

>>>
$$a = [1, 2, 3]$$

>>> $b = a$

```
>>> a = [1, 2, 3]
>>> b = a
>>> b
[1, 2, 3]
>>> a[0] = u'безумный'
```

```
>>> a = [1, 2, 3]
>>> b = a
>>> b
[1, 2, 3]
>>> a[0] = u'безумный'
>>> b
```

```
>>> a = [1, 2, 3]

>>> b = a

>>> b

[1, 2, 3]

>>> a[0] = u'безумный'

>>> b

[u'\u0431\u0435\u0437\u0443\u043c
\u043d\u044b\u0439', 2, 3]
```

```
>>> a = [1, 2, 3]

>>> b = a

>>> b

[1, 2, 3]

>>> a[0] = u'безумный'

>>> b

[u'\u0431\u0435\u0437\u0443\u043c
\u043d\u044b\u0439', 2, 3]
```

```
>>> a = [1, 2, 3]
>>> b = a
>>> b
[1, 2, 3]
>>> b
[1, 2, 3]
>>> a[0] = u'безумный'
>>> b
[u'\u0431\u0435\u0437\u0443\u043c\u0443\u043c\u043d\u044b\u0439', 2, 3]
```

```
>>> a = [1, 2, 3]

>>> b = a

>>> b

[1, 2, 3]

>>> a[0] = u'безумный'

>>> b

[u'\u0431\u0435\u0437\u0443\u043c

\u043d\u044b\u0439', 2, 3]

>>> a = [1, 2, 3]

>>> b

[1, 2, 3]

>>> a[0] = u'безумный'

>>> b
```

```
>>> a = [1, 2, 3]
>>> b = a
>>> b = a[:]
>>> b
[1, 2, 3]
>>> a[0] = u'безумный'
>>> b
[u'\u0431\u0435\u0437\u0443\u043c
\u043d\u044b\u0439', 2, 3]
>>> a = [1, 2, 3]
>>> b = a[:]
>>> b
[1, 2, 3]
>>> b
[1, 2, 3]
>>> b
[1, 2, 3]
>>> b
[1, 2, 3] # unaffected by change to a
```

"Copying" by reference does not actually produce a new object

"Copying" by reference does not actually produce a new object, whereas slicing does.			

"Copying" by reference does not actually produce a new object, whereas slicing does. We can determine whether two objects are the same or not

"Copying" by reference does not actually produce a new object, whereas slicing does. We can determine whether two objects are the same or not, at the moment they are being compared

```
>>> a = [1, 2, 3]
>>> b = a
>>> b
[1, 2, 3]
>>> a[0] = u'безумный'
>>> b
[u'\u0431\u0435\u0437\u0443\u043c
\u043d\u044b\u0439', 2, 3]
>>> id(a) == id(b)
>>> a = [1, 2, 3]
>>> b
= a[:]
>>> b
[1, 2, 3]
>>> b
[1, 2, 3]
>>> b
[1, 2, 3] # unaffected by change to a
>>> id(a) == id(b)
```

```
>>> \alpha = [1, 2, 3]
>>> b = \alpha
>>> b = \alpha
| (1, 2, 3) |
>>> b = \alpha|
| (1, 2, 3) |
>>> a = [1, 2, 3]
>>> b = \alpha|
| (1, 2, 3) |
>>> a = [1, 2, 3]
>>> b = \alpha|
| (1, 2, 3) |
>>> a = [1, 2, 3]
>>> b = \alpha|
| (1, 2, 3) |
>>> a = [1, 2, 3]
>> a = [1, 2, 3]
>>> a = [1, 2, 3]
```

```
>>> a = [1, 2, 3]
>>> b = a[:]
>>> b
>>> 0 = [1, 2, 3]
>>> b = 0
>>> h
[1, 2, 3]
                             >>> a[0] = u'безумный'
>>> a[0] = u'безумный'
>>> h
\u043d\u044b\u0439', 2, 3]
                             >>> id(a) == id(b)
>>> id(a) == id(b)
                             False
True
```

The id() function returns an integer that (in CPython) is the memory address of the argument.

```
>>> a = [1, 2, 3]
>>> b = a[:]
>>> b
>>> 0 = [1, 2, 3]
>>> b = 0
>>> h
[1, 2, 3]
                             >>> a[0] = u'безумный'
>>> a[0] = u'безумный'
>>> h
\u043d\u044b\u0439', 2, 3]
                             >>> id(a) == id(b)
>>> id(a) == id(b)
True
                             False
```

The id() function returns an integer that (in CPython) is the memory address of the argument. Different objects that coexist at some moment have different memory addresses.

>>> import array

```
>>> import array
>>> id(array.array('i', [1, 2, 3])) == id(array.array('i', [1, 2, 3])[:]) # array
```

```
>>> import array
>>> id(array.array('i', [1, 2, 3])) == id(array.array('i', [1, 2, 3])[:]) # array
False
```

```
>>> import array
>>> id(array.array('i', [1, 2, 3])) == id(array.array('i', [1, 2, 3])[:]) # array
False
```

question	list	array
<pre>id(object) == id(object[:])</pre>	False	False

```
>>> import array
>>> id(array.array('i', [1, 2, 3])) == id(array.array('i', [1, 2, 3])[:]) # array
False
```

question	list	array
<pre>id(object) == id(object[:])</pre>	False	False

(Here I use "object" to represent a literal object rather than a variable representing it.)

It isn't surprising that some other Python sequence types display the same behavior as a list when comparing the id() of the original object with a beginning-to-end slice:

question	list	array	tuple
<pre>id(object) == id(object[:])</pre>	False	False	False

It isn't surprising that some other Python sequence types display the same behavior as a list when comparing the id() of the original object with a beginning-to-end slice:

question	list	array	tuple
<pre>id(object) == id(object[:])</pre>	False	False	False

But it may be surprising that not all do...

It isn't surprising that some other Python sequence types display the same behavior as a list when comparing the id() of the original object with a beginning-to-end slice. But it may be surprising that not all do:

It isn't surprising that some other Python sequence types display the same behavior as a list when comparing the id() of the original object with a beginning-to-end slice. But it may be surprising that not all do:

question	list	array	tuple	string
<pre>id(object) == id(object[:])</pre>	False	False	False	True

```
>>> id([1, 2, 3][:]) == id([1, 2, 3][:]) # list
False
>>> id(array.array('i', [1, 2, 3])[:]) == id(array.array('i', [1, 2, 3])[:]) # array
False
```

```
>>> id([1, 2, 3][:]) == id([1, 2, 3][:]) # list
False
>>> id(array.array('i', [1, 2, 3])[:]) == id(array.array('i', [1, 2, 3])[:]) # array
False
>>> id((1, 2, 3)[:]) == id((1, 2, 3)[:]) # tuple
False
```

```
>>> id([1, 2, 3][:]) == id([1, 2, 3][:]) # list
False
>>> id(array.array('i', [1, 2, 3])[:]) == id(array.array('i', [1, 2, 3])[:]) # array
False
>>> id((1, 2, 3)[:]) == id((1, 2, 3)[:]) # tuple
False
>>> id('123'[:]) == id('123'[:]) # string
True
```

```
>>> id([1, 2, 3][:]) == id([1, 2, 3][:]) # list
False
>>> id(array.array('i', [1, 2, 3])[:]) == id(array.array('i', [1, 2, 3])[:]) # array
False
>>> id((1, 2, 3)[:]) == id((1, 2, 3)[:]) # tuple
False
>>> id('123'[:]) == id('123'[:]) # string
True
```

question	list	array	tuple	string
<pre>id(object) == id(object[:])</pre>	False	False	False	True
<pre>id(object[:]) == id(object[:])</pre>	False	False	False	True

```
>>> id([1, 2, 3][:]) == id([1, 2, 3][:]) # list
False
>>> id(array.array('i', [1, 2, 3])[:]) == id(array.array('i', [1, 2, 3])[:]) # array
False
>>> id((1, 2, 3)[:]) == id((1, 2, 3)[:]) # tuple
False
>>> id('123'[:]) == id('123'[:]) # string
True
```

question	list	array	tuple	string
<pre>id(object) == id(object[:])</pre>	False	False	False	True
<pre>id(object[:]) == id(object[:])</pre>	False	False	False	True

Lists and arrays seem always to behave the same way

```
>>> id([1, 2, 3][:]) == id([1, 2, 3][:]) # list
False
>>> id(array.array('i', [1, 2, 3])[:]) == id(array.array('i', [1, 2, 3])[:]) # array
False
>>> id((1, 2, 3)[:]) == id((1, 2, 3)[:]) # tuple
False
>>> id('123'[:]) == id('123'[:]) # string
True
```

question	list/array	tuple	string
<pre>id(object) == id(object[:])</pre>	False	False	True
<pre>id(object[:]) == id(object[:])</pre>	False	False	True

Lists and arrays seem always to behave the same way, but tuples and strings do not.

question	list, array	tuple	string
<pre>id(object) == id(object[:])</pre>	False	False	True
<pre>id(object[:]) == id(object[:])</pre>	False	False	True
var = object id(var) == id(var[:])	False	True	True
var2 = var[:] id(var) == id(var2)	False	True	True

question	list, array	tuple	string
<pre>id(object) == id(object[:])</pre>	False	False	True
<pre>id(object[:]) == id(object[:])</pre>	False	False	True
var = object id(var) == id(var[:])	False	True	True
var2 = var[:] id(var) == id(var2)	False	True	True

question	list, array	tuple	string
<pre>id(object) == id(object[:])</pre>	False	False	True
<pre>id(object[:]) == id(object[:])</pre>	False	False	True
var = object id(var) == id(var[:])	False	True	True
var2 = var[:] id(var) == id(var2)	False	True	True
id(var[:]) == id(var[:])	True	True	True
id(var[:]) == id(var2[:])	True	True	True

question	list, array	tuple	string
<pre>id(object) == id(object[:])</pre>	False	False	True
<pre>id(object[:]) == id(object[:])</pre>	False	False	True
var = object id(var) == id(var[:])	False	True	True
var2 = var[:] id(var) == id(var2)	False	True	True
id(var[:]) == id(var[:])	True	True	True
id(var[:]) == id(var2[:])	True	True	True

ariables and the objects from which they were originally assigned also behave differently:	

Variables and the objects from which they were originally assigned also behave differently:

question	list, array	tuple	string
<pre>id(object) == id(object[:])</pre>	False	False	True
<pre>id(object[:]) == id(object[:])</pre>	False	False	True
var = object id(var) == id(var[:])	False	True	True
var2 = var[:] id(var) == id(var2)	False	True	True
id(var[:]) == id(var[:])	True	True	True
id(var[:]) == id(var2[:])	True	True	True
id(var) == id(object)	False	False	True
id(var2) == id(object)	False	False	True
id(var[:]) == id(object)	True	False	True

Variables and the objects from which they were originally assigned also behave differently:

question	list, array	tuple	string
<pre>id(object) == id(object[:])</pre>	False	False	True
<pre>id(object[:]) == id(object[:])</pre>	False	False	True
var = object id(var) == id(var[:])	False	True	True
var2 = var[:] id(var) == id(var2)	False	True	True
id(var[:]) == id(var[:])	True	True	True
id(var[:]) == id(var2[:])	True	True	True
id(var) == id(object)	False	False	True
id(var2) == id(object)	False	False	True
id(var[:]) == id(object)	True	False	True

>>>
$$id(x[:]) == id([1, 2, 3]) # list$$
 >>> $id(x) == id((1, 2, 3)) # tuple$ False

It seems random.

It seems random. Is it?

It seems random. Is it? Or is there a deep moral correctness in this diverse behavior?				

It seems random. Is it? Or is there a deep moral correctness in this diverse behavior?

In particular, is there some reason why an object being compared with itself should sometimes be considered to be the same object in both cases and sometimes two different objects?

An answer by distribution, not internals: The three implementations CPython, PyPy, and Jython (each v. 2.7) all return different patterns of **True** and **False** with respect to these questions:

question	list, array	tuple	string
<pre>id(object) == id(object[:])</pre>	False	False	C True Py/J
<pre>id(object[:]) == id(object[:])</pre>	False	False	C True Py/J
id(var) == id(var[:])	False	C True Py/J	C/Py True J
id(var) == id(var2)	False	C True Py/J	C/Py True J
id(var[:]) == id(var[:])	C True Py/J	C True Py/J	C/Py True J
id(var[:]) == id(var2[:])	C True Py/J	C True Py/J	C/Py True J
id(var) == id(object)	False	False	C True Py/J
id(var2) == id(object)	False	False	C True Py/J
id(var[:]) == id(object)	C True Py/J	False	C True Py/J

Key: **blue & bold** = **True**; red & non-bold = False

C: CPython; Py: PyPy; J: Jython.

question	list, array	tuple	string
<pre>id(object) == id(object[:])</pre>	False	False	C True Py/J
<pre>id(object[:]) == id(object[:])</pre>	False	False	C True Py/J
id(var) == id(var[:])	False	C True Py/J	C/Py True J
id(var) == id(var2)	False	C True Py/J	C/Py True J
id(var[:]) == id(var[:])	C True Py/J	C True Py/J	C/Py True J
id(var[:]) == id(var2[:])	C True Py/J	C True Py/J	C/Py True J
id(var) == id(object)	False	False	C True Py/J
id(var2) == id(object)	False	False	C True Py/J
id(var[:]) == id(object)	C True Py/J	False	C True Py/J

That is the main point of this presentation.

question	list, array	tuple	string
<pre>id(object) == id(object[:])</pre>	False	False	C True Py/J
<pre>id(object[:]) == id(object[:])</pre>	False	False	C True Py/J
id(var) == id(var[:])	False	C True Py/J	C/Py True J
id(var) == id(var2)	False	C True Py/J	C/Py True J
id(var[:]) == id(var[:])	C True Py/J	C True Py/J	C/Py True J
id(var[:]) == id(var2[:])	C True Py/J	C True Py/J	C/Py True J
id(var) == id(object)	False	False	C True Py/J
id(var2) == id(object)	False	False	C True Py/J
id(var[:]) == id(object)	C True Py/J	False	C True Py/J

That is the main point of this presentation. (Note that Jython tests False everywhere that CPython tests True here; PyPy is mixed.)

question	list, array	tuple	string
<pre>id(object) == id(object[:])</pre>	False	False	C True Py/J
<pre>id(object[:]) == id(object[:])</pre>	False	False	C True Py/J
id(var) == id(var[:])	False	C True Py/J	C/Py True J
id(var) == id(var2)	False	C True Py/J	C/Py True J
id(var[:]) == id(var[:])	C True Py/J	C True Py/J	C/Py True J
id(var[:]) == id(var2[:])	C True Py/J	C True Py/J	C/Py True J
id(var) == id(object)	False	False	C True Py/J
id(var2) == id(object)	False	False	C True Py/J
id(var[:]) == id(object)	C True Py/J	False	C True Py/J

That is the main point of this presentation. (Note that Jython tests False everywhere that CPython tests True here; PyPy is mixed.) In sum: the behavior of the id() function is uniform neither with respect to the various sequences nor among the three main implementations.

question	list, array	tuple	string
<pre>id(object) == id(object[:])</pre>	False	False	C True Py/J
<pre>id(object[:]) == id(object[:])</pre>	False	False	C True Py/J
id(var) == id(var[:])	False	C True Py/J	C/Py True J
id(var) == id(var2)	False	C True Py/J	C/Py True J
id(var[:]) == id(var[:])	C True Py/J	C True Py/J	C/Py True J
id(var[:]) == id(var2[:])	C True Py/J	C True Py/J	C/Py True J
id(var) == id(object)	False	False	C True Py/J
id(var2) == id(object)	False	False	C True Py/J
id(var[:]) == id(object)	C True Py/J	False	C True Py/J

That is the main point of this presentation. (Note that Jython tests False everywhere that CPython tests True here; PyPy is mixed.) In sum: the behavior of the id() function is uniform neither with respect to the various sequences nor among the three main implementations. It seems doubtful that this is a matter of performance following prescription.

(For reference here is how "identity" is defined in the three versions illustrated here:

CPython: "Return the 'identity' of an object. This is an integer which is guaranteed to be unique and constant for this object during its lifetime. Two objects with non-overlapping lifetimes may have the same id() value. CPython implementation detail: This is the address of the object in memory." https://docs.python.org/2.7/library/functions.html?#id. Python 2.7.8 (default, Jul 2 2014, 10:14:46) [GCC 4.2.1 Compatible Apple LLVM 5.1 (clang-503.0.40)] on darwin

PyPy: "Using the default GC (called minimark), the built-in function id() [of PyPy] works like it does in CPython. With other GCs it returns numbers that are not real addresses (because an object can move around several times) and calling it a lot can lead to performance problem." http://pypy.readthedocs.org/en/latest/cpython_differences.html Python 2.7.6 (32f35069a16d, Jun 06 2014, 20:12:47)

Jython: "Return the 'identity' of an object. This is an integer (or long integer) which is guaranteed to be unique and constant for this object during its lifetime. Two objects with non-overlapping lifetimes may have the same id() value. (Implementation note: this is the address of the object.)" http://www.jython.org/docs/library/functions.html [PyPy 2.3.1 with GCC 4.2.1 Compatible Apple LLVM 5.0 (clang-500.2.79)] on darwin Jython 2.7b2 (default:a5bc0032cf79+, Apr 22 2014, 21:20:17) [Java HotSpot(TM) 64-Bit Server VM (Oracle Corporation)] on java1.7.0_51)

>>> def test_list():	>>> def test_list():
$x = [1, 2, 3][:]$	print id([1, 2, 3][:])
print id(x[:])	print id([1, 2, 3][:])
print id(x[:])	print id([1, 2, 3][:])
print id(x[:])	print id([1, 2, 3][:])
print id(x[:])	
>>> test_list()	>>> test_list()
4451744728	4451745160
4451744728	4451676816
4451744728	4451745160
4451744728	4451676816

>>> def test_list():	>>> def test_list():
$x = [1, 2, 3][:]$	print id([1, 2, 3][:])
print id(x[:])	print id([1, 2, 3][:])
print id(x[:])	print id([1, 2, 3][:])
print id(x[:])	print id([1, 2, 3][:])
print id(x[:])	
>>> test_list()	>>> test_list()
4451744728	4451745160
4451744728	4451676816
4451744728	4451745160
4451744728	4451676816

For a list, the literal object uses two alternating memory addresses in this example, while a variable uses the same memory address.

>>> def test_list():	>>> def test_list():
$x = [1, 2, 3][:]$	print id([1, 2, 3][:]) # do this four times
print id(x[:]) # do this four times	
>>> test_list()	>>> test_list()
4451744728	4451745160
4451744728	4451676816

The three implementations behave differently in this respect;

>>> def test_list():	>>> def test_list():
$x = [1, 2, 3][:]$	print id([1, 2, 3][:]) # do this four times
print id(x[:]) # do this four times	
>>> test_list()	>>> test_list()
4451744728	4451745160
4451744728	4451676816

The three implementations behave differently in this respect; Jython again is always False:

>>> def test_list():	>>> def test_list():
$x = [1, 2, 3][:]$	print id([1, 2, 3][:]) # do this four times
print id(x[:]) # do this four times	
>>> test_list()	>>> test_list()
4451744728	4451745160
4451744728	4451676816

The three implementations behave differently in this respect; Jython again is always False:

question	list, array	tuple	string
id(var[:]), 4x in fn same	C True Py/J	C True Py/J	C/Py True J
id(object[:]), 4x in fn same	C False*; Py/J	False	C/Py True J

 $^{^\}star$ IDs appear in alternation.

劇終