Subject: Re: Where vs Histogram vs ??
Posted by Pavel A. Romashkin on Wed, 23 Oct 2002 17:29:45 GMT
View Forum Message <> Reply to Message

Hi Andrew,
Sorry for delaying the answer.
No, no, no. No. It needs to be just what it is. It will be an array of just 9 pointers. Each of them points to a vector (well, except for the last one which is a matrix), and as such is searchable quite quickly using WHERE.
You may notice that for an array of structures:

```
a = {a: 0, b: 0.0, c: fltarr(5)}
a = replicate(a, 1000)
help, a.(0)
;<Expression>    INT       = Array[1000]
help, a.(2)
;<Expression>    FLOAT     = Array[5, 1000]
```

Therefore, when you loop over just *fields* of a structure array, you get the contents of the entire array. In your case, this is perfect for indexing the data. I use this a lot - it allows to shift arrays throughout the entire structure array just as if it were a plain matrix or vector, and is just as fast.
As I said, you can basically do away with the sreucture array becasue now your 9-element pointer array contains everything the old structure array contained. In fact, yopu can dump the old array do free up some RAM, but that is not critical. Also, in a general case, you want only to include those fields in the ptr array that you use for searching, and then use the resulting index to extract the data from the original structure array.
Regarding memory use:

```
; Here, A is an array of structures of exactly your type of size 16 mln.
; I have nothing else in the IDL session.
IDL> help, /mem
heap memory used:  512482366, max:  512483544, gets:     1719, frees:
 1167
IDL> ind = ptrarr(n_tags(a)
IDL> for i = 0, n_tags(a)-1 do ind[i] = ptr_new(a.(i))
; The above takes less than a minute
IDL> help, /mem
heap memory used: 1024484012, max: 1024484732, gets:     3656, frees:
 3093
```

As expected, the memory use doubles; if that's a problem, discard the original array.

Hope this helps.
Pavel

Andrew Cool wrote:
> Should this be something like
>
> ind = ptrarr(N_Tags(data_st) * 15425228L)
>
> given that N_Tags(data_st) only returns a value of 9, which concurs
> with Tag_Names(data_st), such that we effectively have
>
> ind = ptrarr(9 * 15425228L)
>
>
> Now that's a scary sized ptrarr.
>
> Given that you say :-
>
>> On my machine the RAM used by both structure and pointer index barely
>> reaches 1010 Mb, so I have room for further calculations.
>
> and assuming you've used the figure of 15425228, then I obviously
> don't
> understand your example... ;-)
>
> Would you mind elaborating a bit, in words of one brain cell or less?
>
> Thanks,
>
> Andrew