
Subject: Re: Some questions of efficiency when matching items in lists

Posted by [Craig Markwardt](#) on Fri, 01 Feb 2013 00:59:16 GMT

[View Forum Message](#) <> [Reply to Message](#)

On Thursday, January 31, 2013 7:05:40 PM UTC-5, Bogdanovist wrote:

> I have a couple of questions about how to efficiently match items in lists. There are two operations that are done many thousands of times in my processing and are causing a bottleneck. Even small improvements would be welcome.

For your first question, the obvious thing is that you are reading and writing a file for every operation. File I/O is slow. If you can, keep your values in memory and only write them out when necessary. At the very least, only rewrite your file when you have to; otherwise just /APPEND to the end.

For your second question: if DATA_ADD has a small number of elements then you are probably doing the best that you can do. You might check out MATCH or MATCH2 in the IDL astronomy library, which have been optimized for cross-matching a lot of elements. Another possibility is to create a hash table indexed by time; this has the benefit of rapid access, but you lose the ability to perform vector operations upon the whole table.

Craig
