

UrlManager

Serge Emond

Copyright © Copyright1996 by Serge Emond

COLLABORATORS

	<i>TITLE :</i> UrlManager	
<i>ACTION</i>	<i>NAME</i>	<i>DATE</i>
WRITTEN BY	Serge Emond	February 12, 2023
<i>SIGNATURE</i>		

REVISION HISTORY

NUMBER	DATE	DESCRIPTION	NAME

Contents

1	UrlManager	1
1.1	UrlManager.guide/Contents	1
1.2	UrlManager.guide/About English	1
1.3	UrlManager.guide/Introduction	2
1.4	UrlManager.guide/Features	2
1.5	UrlManager.guide/Legal information	2
1.6	UrlManager.guide/License	3
1.7	UrlManager.guide/No warranty	3
1.8	UrlManager.guide/Requirements	3
1.9	UrlManager.guide/History	4
1.10	UrlManager.guide/What's next	4
1.11	UrlManager.guide/Bugs	4
1.12	UrlManager.guide/Credits	4
1.13	UrlManager.guide/Author	4
1.14	UrlManager.guide/This is BETA	4
1.15	UrlManager.guide/Arguments	5
1.16	UrlManager.guide/Arguments/PORT	5
1.17	UrlManager.guide/Types	5
1.18	UrlManager.guide/Depth	5
1.19	UrlManager.guide/IDs	5
1.20	UrlManager.guide/Using ARexx	6
1.21	UrlManager.guide/ARexx/AddUrl	6
1.22	UrlManager.guide/ARexx/FlushAll	7
1.23	UrlManager.guide/ARexx/GetDepth	7
1.24	UrlManager.guide/ARexx/GetHid	8
1.25	UrlManager.guide/ARexx/GetInMem	8
1.26	UrlManager.guide/ARexx/GetType	8
1.27	UrlManager.guide/ARexx/GetUrl	9
1.28	UrlManager.guide/ARexx/KillPattern	9
1.29	UrlManager.guide/ARexx/KillType	9

1.30	UrlManager.guide/ARexx/KillUrl	10
1.31	UrlManager.guide/ARexx/LoadFile	10
1.32	UrlManager.guide/ARexx/MatchName	10
1.33	UrlManager.guide/ARexx/Quit	11
1.34	UrlManager.guide/ARexx/ReadFile	11
1.35	UrlManager.guide/ARexx/SaveFile	12
1.36	UrlManager.guide/ARexx/Search	12
1.37	UrlManager.guide/ARexx/SetDepth	13
1.38	UrlManager.guide/ARexx/SetIOSize	13
1.39	UrlManager.guide/ARexx/SetType	13
1.40	UrlManager.guide/ARexx/SetUrl	14

Chapter 1

UrlManager

1.1 UrlManager.guide/Contents

UrlManager

Version 0.1

Copyright (c) 1996 Serge Emond

[Important note about this English documentation](#)

[Introduction](#)

[Features](#)

[Legal information](#)

[Requirements](#)

[This is a beta release](#)

[Usage](#)

[ARexx](#)

[Known bugs](#)

[The future](#)

[Credits](#)

[Author](#)

[Program history](#)

1.2 UrlManager.guide/About English

A Note About English

You will find in this document sentences incorrectly formed. Bad English. Very bad. My orthograph is bad, my verbs are bad, everything is bad.

I write the documentation of my programs in English because it's the only way to write ONE document that anyone can understand.

Sometimes when I browse new files on AmiNet, I see files that I must grab... I grab them. It always happens to the one that seems the most interesting: After minutes of downloading, no English documentation. Even the program is in a foreign language. The first thing I do after such a discovery is to delete the lot. (Except if the foreign language is French, but it never happened to me.)

It would be *LOT* faster for me to write this document in French but...

Anyway, if you didn't get the point, all the flafa above means:

There are many grammatical errors in this document and I don't care about them. What is important to me for now is to produce a readable documentation, nothing more. (And I'm sure it's readable...)

1.3 UrlManager.guide/Introduction

Introduction

You enter on Yahoo, do a search about "Brain" and find a marvelous site containing lots of data on the Brain. You want to grab everything but all the information is scattered across thousands of .html files. A real nightmare to grab everything with a web browser.

More, one of your hobbies is to view pictures. It relaxes yourself and you have enough space on your harddrive not to bother with selecting the images before you download them. Yesterday you found a superb site containing thousands of sunsets. You don't want to stay 20 hours to click each one and press enter in the requester. Worst some of them are not clickable and your browser don't have any option to save those to disk, you have to find the original filename in the html file and copy the file from the disk cache (if there is one!).

The solution: a robot that can download a file, parse it to find other files, grab them, and so on. More, it should grab some files and forget others. It should allow to download only during x hours, no more than y megs or z files.

UrlManager will open an ARexx port and wait for commands to add, remove or search urls.

1.4 UrlManager.guide/Features

Features

- Written in C, which means QUICK compared with ARexx scripts
- Each url is checked to see if already present in the list
- Each url has a "type" which can be any uppercase character
- Each url has a "depth" which is a long integer that can be used to indicate how "deep" to search this url for other urls.
- Many ways to search an url in the list, including AmigaDOS patterns
- Can read lists of URLs from file
- Can save an url-list containing all data (depth, type)
- And many more :)

1.5 UrlManager.guide/Legal information

Legal information

License

No warranty

1.6 UrlManager.guide/License

License

UrlManager is released under the concept of freeware. This means you are allowed to use and copy this program freely, as long as the following requirements are fulfilled:

All files are copied without any alterations or modifications. If any extra files are added, it must be obvious that they do not belong to the original distribution, and that they do not need to be included in any redistribution. Exception: So called "BBS ads" may not be added.

The copying is done on a non-commercial basis. A small fee to cover media costs etc. may be charged.

The copier is not claiming the copyright of this program.

Any exceptions from the above require a written permission from the author.

If you want to publish this program on a cover disk or similar, contact me first for approval (to make sure you have the latest version etc). I then expect a copy of the issue in question in return (additional contributions are welcomed :).

1.7 UrlManager.guide/No warranty

No warranty

There is no warranty for the programs, to the extent permitted by applicable law. Except when otherwise stated in writing the copyright holder and/or other parties provide the programs "as is" without warranty of any kind, either expressed or implied, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose. The entire risk as to the quality and performance of the programs is with you. Should the programs prove defective, you assume the cost of all necessary servicing, repair or correction.

In no event unless required by applicable law or agreed to in writing will any copyright holder, or any other party who may redistribute the programs as permitted above, be liable to you for damages, including any general, special, incidental or consequential damages arising out of the use or inability to use the programs (including but not limited to loss of data or data being rendered inaccurate or losses sustained by you or third parties or a failure of the programs to operate with any other programs), even if such holder or other party has been advised of the possibility of such damages.

1.8 UrlManager.guide/Requirements

Requirements

UrlManager requires at least OS v37 (2.0 release). You must know however that if used on a pre-v39 (3.0) machine, there is a bug in dos.library: character-classes ([x-y]) are not converted properly to uppercase. UrlManager won't uppercase them either. I suppose it will work OK if you use uppercase [x-y] patterns. (See you AmigaDOS manual or dos.lib's autodocs about patterns)

1.9 UrlManager.guide/History

The Past

Version 0 Revision 1 - 16.8.1996 and before

Nothing recorded.

1.10 UrlManager.guide/What's next

What's Next - The To Do List

- Multiple lists support
- Add a user field
- AddUrl: return ID of existing url if RC=1.

1.11 UrlManager.guide/Bugs

Known Bugs

- Patterns, see [Requirements](#) .

1.12 UrlManager.guide/Credits

Credits

UrlManager was entirely written in 'C' by [Serge Emond](#) .

1.13 UrlManager.guide/Author

About the author

If you wish to send me comments, bug reports, gifts, money or whatever, then use this address:

Serge Emond

3392 des Anémones

Jonquière, Québec

Canada, G7X 7V3

I do not live there most of the time but whatever you send WILL reach me.

(!) Currently no email (!)

1.14 UrlManager.guide/This is BETA

This is a BETA release

This release is BETA. This means it has not been fully tested. I release it only because it seems stable and it works well on my computer. As stated in [legal issues](#) I may not be responsible for any damage caused by the use of this program.

1.15 UrlManager.guide/Arguments

Arguments

UrlManager can only be used from a Shell.

You can break it at anytime by sending it a CTRL-C break.

Please note that UrlManager is pure and can be made resident.

Text arguments

PORT Name of the ARExx port to open.

1.16 UrlManager.guide/Arguments/PORT

PORT

Required. Name of the ARExx port to open. You should use uppercase only if you plan not to always enquote the port-name in your arexx scripts.

1.17 UrlManager.guide/Types

Types

A type in UrlManager is a single character which can be any printable character. (ie letters and numbers)

If it is a letter, the type is uppercased. This way, in your rexx scripts, 'a' and 'A' are always the same thing.

Examples:

Q Queued

R Received

F Failed

1.18 UrlManager.guide/Depth

Depth

For UrlManager, a depth is simply a signed long integer.

For an url grabber, it can be used to indicate the level of recursion of a single url. For example a depth >0 could mean "Scan me and download urls contained in me". All urls contained in this one are added with a depth of 1, and then a depth of 0, which means "don't scan me".

When no depth is specified when adding an url, either with the **AddUrl** ARExx command or by loading an url list, the depth is set to "-1".

1.19 UrlManager.guide/IDs

IDs

Each url has a unique ID in UrlManager. This ID is used to get information on a specific url.

1.20 UrlManager.guide/Using ARexx

Using ARexx with UrlManager

When started, UrlManager does ONE thing: wait for arexx messages to come.

Each time a command is received, the RC variable and a variable named "UMRES" are both set.

Any command can return at least one of these codes in RC:

10 Command unknown/mispelled

15 Error setting secondary result variable (UMRES)

Here's a list of all ARexx commands known of UrlManager:

ADDURL

FLUSHALL

GETDEPTH

GETHID

GETINMEM

GETTYPE

GETURL

KILLPATTERN

KILLTYPE

KILLURL

LOADFILE

MATCHNAME

QUIT

READFILE

SAVEFILE

SEARCH

SETDEPTH

SETIOSIZE

SETTYPE

SETURL

1.21 UrlManager.guide/ARexx/AddUrl

AddUrl TYPE/A,URL/A,DEPTH/N

AddUrl adds an url to UM's url list.

Arguments:

TYPE/A:

A single uppercase character representing the type Link TYPES } of the url.

URL/A:

The url itself :)

DEPTH/N:

The **depth** of the url. If omitted, -1 is used.

RC:

0 Added. Url's ID is in UMRES.

1 The url was already in the list.

2 Error adding url. (ie not enough memory)

5 Error parsing args.

UMRES:

ID of the url.

1.22 UrlManager.guide/ARexx/FlushAll

FlushAll

Flush the entire list and reset the **IDs** .

Arguments:

None.

RC:

0 Ok.

99 Could not create new memory pools.

You really should issue a **Quit** command
if you receive this.

UMRES:

Nothing.

1.23 UrlManager.guide/ARexx/GetDepth

GetDepth ID/A/N

Get the depth of an url.

Arguments:

ID/A/N:

The **ID** of the url to get depth from.

RC:

0 Depth is in UMRES.

1 This ID is not in use.

5 Error parsing args.

UMRES:

Depth of the url.

1.24 UrlManager.guide/ARexx/GetHid

GetHID

Get the highest ID number.

Arguments:

None.

RC:

0 Ok.

UMRES:

Highest ID number.

1.25 UrlManager.guide/ARexx/GetInMem

GetInMem

Get the total number of URLs in the list.

Arguments:

None.

RC:

0 Ok.

UMRES:

Total number of urls in the list.

1.26 UrlManager.guide/ARexx/GetType

GetType ID/A/N

Get the type of an url.

Arguments:

ID/A/N:

The ID of the url to get type from.

RC:

0 Ok.

1 ID not in use.

5 Error parsing args.

UMRES:

The type of the url requested.

1.27 UrlManager.guide/ARexx/GetUrl

GetUrl ID/A/N

Get the name of an url.

Arguments:

ID/A/N:

ID of the url to get name from.

RC:

0 Ok. Name in UMRES.

1 ID not in use.

5 Error parsing args.

UMRES:

Url.

1.28 UrlManager.guide/ARexx/KillPattern

KillPattern PATTERN/A

Kill all urls matching an AmigaDOS pattern. (Or an exact url name if no pattern-chars are given :)

Arguments:

PATTERN/A:

An AmigaDOS pattern.

RC:

0 Ok.

5 Error parsing args.

UMRES:

Quantity of url killed. 0 can either mean that the pattern could not be parsed or that there were no url matching pattern.

1.29 UrlManager.guide/ARexx/KillType

KillType TYPE/A

Kill all urls of a specific **type** .

Arguments:

TYPE/A:

A single uppercase character representing the **type** .

RC:

0 Ok.

5 Error parsing args.

UMRES:

Quantity of url killed.

1.30 UrlManager.guide/ARexx/KillUrl

KillUrl ID/A/N

Kill a specific url.

Arguments:

ID/A/N:

Url **ID** to kill.

RC:

0 Killed.

1 ID not in use.

5 Error parsing args.

UMRES:

Nothing.

1.31 UrlManager.guide/ARexx/LoadFile

LoadFile FILE/A

Read a file containing an url list. One url/line, each line like this:

<Type> <Depth> <Url>

Only spaces (DEC 32, HEX 20, " ") are allowed between the 3 fields. (NO tabs...)

Arguments:

FILE/A:

Filename.

RC:

0 Everything went OK. QUantity of url added in UMRES.

1 Error opening/parsing file.

5 Error parsing args.

UMRES:

Quantity of url added.

See also:

SAVEFILE

1.32 UrlManager.guide/ARexx/MatchName

MatchName NAME/A,NOCASE/S

Find an url with a specific name.

Arguments:

NAME/A:

Name of the url.

NOCASE/S:

If specified, do a case insensitive search.

RC:

0 Ok. ID in UMRES.

1 No match.

5 Error parsing args.

UMRES:

ID of the url.

1.33 UrlManager.guide/ARexx/Quit

Quit

Tells UrlManager to free its memory and to exit.

Arguments:

None.

RC:

0 I'm quitting.

UMRES:

Nothing.

1.34 UrlManager.guide/ARexx/ReadFile

ReadFile FILE/A,TYPE/A,DEPTH/N

Read a file containing one url per line. Each line must only contain the url.

Arguments:

FILE/A:

The name of the file to read.

TYPE/A:

The type Link TYPES} each url added will have.

DEPTH/N:

Depth each url will have. Default: -1.

RC:

0 Ok. Quantity of url added in UMRES.

1 Error opening/parsing file.

5 Error parsing args.

UMRES:

Quantity of url added.

See also:

SaveFile

1.35 UrlManager.guide/ARexx/SaveFile

SaveFile FILE/A,TYPE/K,PATTERN/K,APPEND/S,FULL/S

Selectively save url(s) to disk, either in **ReadFile** 's or **LoadFile** 's format.

Arguments:

FILE/A:

Name of to file to write urls to.

TYPE/K:

If specified, save only urls of that type.

PATTERN/K:

If specified, save only urls matching that AmigaDOS pattern.

APPEND/S:

If specified, append to file instead of overwriting.

FULL/S:

If present, save a **LoadFile** file, if absent, save a **ReadFile** file.

RC:

0 Ok.

1 Error saving file.

5 Error parsing args.

UMRES:

Nothing.

1.36 UrlManager.guide/ARexx/Search

Search MINID/A/N,PATTERN/K,TYPE/K

Search the list (lowest ID first) for an url.

Arguments:

MINID/A/N

Required. Lowest ID number to scan. 0 scans all of them..

PATTERN/K:

AmigaDOS pattentr that must be matched if specified.

TYPE/K:

If specified, scan only that **type** .

RC:

0 Ok.

5 Error parsing args.

UMRES:

ID of the url or 0 if nothing found.

1.37 UrlManager.guide/ARexx/SetDepth

SetDepth ID/A/N,DEPTH/A/N

Change the **depth** of an url.

Arguments:

ID/A/N:

ID of the url.

DEPTH/A/N:

New **depth** of the url.

RC:

0 Ok.

1 ID not in use.

5 Error parsing args.

UMRES:

Nothing.

1.38 UrlManager.guide/ARexx/SetIOSize

SetIOSize SIZE/A/N

Set the size of input/output buffer.

Arguments:

SIZE/A/N:

Size in bytes. If < 1024, 1024 is used. Default: 4096.

RC:

0 Ok.

5 Error parsing args.

UMRES:

Nothing.

1.39 UrlManager.guide/ARexx/SetType

SetType ID/A/N,TYPE/A

Change the **type** of an url.

Arguments:

ID/A/N:

ID of the url.

TYPE/A:

New **type** of the url.

RC:

0 Ok.

1 ID not in use.

5 Error parsing args.

UMRES:

Nothing.

1.40 UrlManager.guide/ARexx/SetUrl

SetDepth ID/A/N,URL/A

Change the name of an url.

Arguments:

ID/A/N:

ID of the url.

URL/A:

New name of the url.

RC:

0 Ok.

1 ID not in use.

5 Error parsing args.

UMRES:

Nothing.
