Subject: Re: str_sep bug
Posted by Phil Williams on Mon, 25 Nov 1996 08:00:00 GMT
View Forum Message <> Reply to Message

Sorry to answer my own question. BUT, after posting I searched the str_sep routine and think I located the bug. It's at line 122 of str_sep.pro. The line reads:

else arr(i) = strmid(s, spos < strlen(spos) - 1, strlen(s))

but should read:

else arr(i) = strmid(s, spos-strlen(sep), strlen(s))

I still think that there is a "bug." Although some would call it a feature? It's that the last element in the seperated list still has the seperator attached.

Hope this helps,
Phil

Phil Williams wrote:
>
> I think that I have found a bug in the str_sep.pro function. Observe
> the following:
>
> IDL> t = 'apples, oranges, peaches'
> IDL> print,str_sep(t,',')
> apples oranges , peaches
> IDL> print,str_sep(t,', ')
> apples oranges nges, peaches
>
> It appears that the last element is messed up when you add the space to
> the separator it fails on the last element.
>
> Any help would be appreciated.
>
> Phil
> --
> /** ********/
> Phil Williams, Ph.D.
> Research Instructor
> Children's Hospital Medical Center "One man gathers what
> Imaging Research Center another man spills..."
> 3333 Burnet Ave. -The Grateful Dead
> Cincinnati, OH 45229
> email: williams@irc.chmcc.org
> URL: http://scuttle.chmcc.org/~williams/

> /** ********/

--
/** ********/
Phil Williams, Ph.D.
Research Instructor
Children's Hospital Medical Center "One man gathers what
Imaging Research Center another man spills..."
3333 Burnet Ave. -The Grateful Dead
Cincinnati, OH 45229
email: williams@irc.chmcc.org
URL: http://scuttle.chmcc.org/~williams/
/** ********/

Subject: Re: str_sep bug
Posted by thompson on Tue, 26 Nov 1996 08:00:00 GMT
View Forum Message <> Reply to Message

Phil Williams <williams@irc.chmcc.org> writes:

> I am using v1.9 of str_sep dated December 1995.

Evidently there was a bug introduced somewhere between version 1.3 and 1.9.
Maybe somebody from RSI could comment. In the meantime, here's version 1.3.

Bill Thompson

```
 ============================================================
==================
; $Id: str_sep.pro,v 1.3 1995/01/06 21:59:22 dave Exp $
; Copyright (c) 1992-1995, CreaSo Creative Software Systems GmbH,
; and Research Systems, Inc.  All rights reserved.
; Unauthorized reproduction prohibited.
;+
; NAME:
;    STR_SEP
;
; PURPOSE:
;    This routine cuts a string into pieces which are separated by the
;    separator string.
; CATEGORY:
;    String processing.
; CALLING SEQUENCE:
;    arr = STR_SEP(string, separator)
;
; INPUTS:
;    str - The string to be separated.
```

```
;    sep - The separator.
;
; KEYWORDS:
;    ESC = escape character.  Only valid if separator is a single character.
;  Characters following the escape character are treated
;  literally and not interpreted as separators.
;  For example, if the separator is a comma,
;  and the escape character is a backslash, the character
;  sequence 'a\,b' is a single field containing the characters
;  'a,b'.
;    REMOVE = if set, remove all blanks from fields.
;    TRIM = if set, remove only leading and trailing blanks from fields.
;
; OUTPUT:
;    An array of strings as function value.
;
; COMMON BLOCKS:
;    None
;
; SIDE EFFECTS:
;    No known side effects.
;
; RESTRICTIONS:
;    None.
;
; EXAMPLE:
;    array = STR_SEP ("ulib.usca.test", ".")
;
; MODIFICATION HISTORY:
; July 1992, AH, CreaSo  Created.
; December, 1994, DMS, RSI Added TRIM and REMOVE.
;-
function STR_SEP, s, sep, REMOVE = remove, TRIM = trim, ESC=esc


spos = 0L
if n_elements(esc) gt 0 then begin  ;Check for escape character?
  if strpos(s, esc) lt 0 then goto, no_esc ;None in string, use fast case
  besc = (byte(esc))(0)
  bsep = (byte(sep))(0)
  new = bytarr(strlen(s)+1)
  new(0) = byte(s)
  j = 0
  for i=0, n_elements(new)-2 do begin
    if new(i) eq besc then begin
 new(j) = new(i+1)
 i = i + 1
    endif else if new(i) eq bsep then new(j) = 1b $   ;Change seps to 1b char
```

```
    else new(j) = new(i)
    j = j + 1
    endfor
  new = string(new(0:j-1))
  w = where(byte(new) eq 1b, count)  ;where seps are...
  arr = strarr(count+1)
  for i=0, count-1 do begin
 arr(i) = strmid(new, spos, w(i)-spos)
 spos = w(i) + 1
 endfor
  arr(count) = strmid(new, spos, strlen(s))  ;Last element
  goto, done
  endif   ;esc

no_esc:
if strlen(sep) eq 1 then begin ;Single character separator?
    w = where(byte(s) eq (byte(sep))(0), count)  ;where seps are...
    arr = strarr(count+1)
    for i=0, count-1 do begin
 arr(i) = strmid(s, spos, w(i)-spos)
 spos = w(i) + 1
 endfor
    arr(count) = strmid(s, spos, strlen(s))  ;Last element
endif else begin  ;Multi character separator....
    n = 0     ; Determine number of seperators in string.
    repeat begin
 pos = strpos (s, sep, spos)
 spos = pos + strlen(sep)
 n = n+1
    endrep until pos eq -1

    arr = strarr(n)    ; Create result array
    spos = 0
    for i=0, n-1 do begin   ; Separate substrings
      pos = strpos (s, sep, spos)
      if pos ge 0 then arr(i) = strmid (s, spos, pos-spos) $
      else arr(i) = strmid(s, spos, strlen(s))
      spos = pos+strlen(sep)
   endfor
endelse

done:
if keyword_set(trim) then arr = strtrim(arr,2) $
else if keyword_set(remove) then arr = strcompress(arr, /REMOVE_ALL)
return, arr
end
```

Subject: Re: str_sep bug
Posted by Phil Williams on Tue, 26 Nov 1996 08:00:00 GMT
View Forum Message <> Reply to Message

I am using v1.9 of str_sep dated December 1995.

--
/** ********/
Phil Williams, Ph.D.
Research Instructor
Children's Hospital Medical Center "One man gathers what
Imaging Research Center another man spills..."
3333 Burnet Ave. -The Grateful Dead
Cincinnati, OH 45229
email: williams@irc.chmcc.org
URL: http://scuttle.chmcc.org/~williams/
/** ********/

Subject: Re: str_sep bug
Posted by thompson on Tue, 26 Nov 1996 08:00:00 GMT
View Forum Message <> Reply to Message

Phil Williams <williams@irc.chmcc.org> writes:

> Sorry to answer my own question. BUT, after posting I searched the
> str_sep routine and think I located the bug. It's at line 122 of
> str_sep.pro. The line reads:

Sorry to have forgotten this point in the previous message, but one of the
things that leads me to believe you have a nonstandard version is that both of
the versions I've looked at are shorter than 122 lines.

Bill Thompson

Subject: Re: str_sep bug
Posted by thompson on Tue, 26 Nov 1996 08:00:00 GMT
View Forum Message <> Reply to Message

Phil Williams <williams@irc.chmcc.org> writes:

> Sorry to answer my own question. BUT, after posting I searched the
> str_sep routine and think I located the bug. It's at line 122 of
> str_sep.pro. The line reads:

> else arr(i) = strmid(s, spos < strlen(spos) - 1, strlen(s))

> but should read:

> else arr(i) = strmid(s, spos-strlen(sep), strlen(s))

> I still think that there is a "bug." Although some would call it a
> feature? It's that the last element in the seperated list still has the
> seperator attached.

 (rest deleted)

What version of str_sep.pro are you using? I cannot replicate your problem with either the version (1.1) which comes with IDL 3.6 or the version (1.3) which comes with IDL 4.0.1. My suspicion is that you're using some non-standard version of str_sep.pro. I would certainly consider the behavior you describe as a serious bug.

Bill Thompson
