

Ruby master - Feature #836

Patches for StringScanner, adding #size, #captures and #values_at

12/08/2008 11:27 AM - apeiros (Stefan Rusterholz)

| | |
|--|-------------------------|
| Status: | Closed |
| Priority: | Normal |
| Assignee: | nobu (Nobuyoshi Nakada) |
| Target version: | 2.6 |
| Description =begin The methods are named to be consistent with MatchData. My C-fu isn't very strong, so a review of the patch would be appreciated. I attached the revised code and a diff file. If other formats are wished, please tell. If I need to submit something else with that patch, please tell. Regards Stefan =end | |

Associated revisions

Revision 256c8886 - 11/29/2017 07:57 AM - nobu (Nobuyoshi Nakada)

strscan.c: add MatchData-like methods

- ext/strscan/strscan.c: added size, captures and values_at to StringScanner, shorthands of accessing the matched data. based on the patch by apeiros (Stefan Rusterholz) at [ruby-core:20412]. [Feature #836]

git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/trunk@60929 b2dd03c8-39d4-4d8f-98ff-823fe69b080e

Revision 60929 - 11/29/2017 07:57 AM - nobu (Nobuyoshi Nakada)

strscan.c: add MatchData-like methods

- ext/strscan/strscan.c: added size, captures and values_at to StringScanner, shorthands of accessing the matched data. based on the patch by apeiros (Stefan Rusterholz) at [ruby-core:20412]. [Feature #836]

Revision 60929 - 11/29/2017 07:57 AM - nobu (Nobuyoshi Nakada)

strscan.c: add MatchData-like methods

- ext/strscan/strscan.c: added size, captures and values_at to StringScanner, shorthands of accessing the matched data. based on the patch by apeiros (Stefan Rusterholz) at [ruby-core:20412]. [Feature #836]

Revision 60929 - 11/29/2017 07:57 AM - nobu (Nobuyoshi Nakada)

strscan.c: add MatchData-like methods

- ext/strscan/strscan.c: added size, captures and values_at to StringScanner, shorthands of accessing the matched data. based on the patch by apeiros (Stefan Rusterholz) at [ruby-core:20412]. [Feature #836]

History

#1 - 12/09/2008 02:34 AM - nobu (Nobuyoshi Nakada)

Hi,

At Mon, 8 Dec 2008 11:20:55 +0900,

Stefan Rusterholz wrote in [ruby-core:20412]:

The methods are named to be consistent with MatchData.
My C-fu isn't very strong, so a review of the patch would be appreciated.
I attached the revised code and a diff file. If other formats are wished, please tell.
If I need to submit something else with that patch, please tell.

- Variable declaration is disallowed after executing statements.
- call-seq: for captures was `size`.

Index: ext/strscan/strscan.c

```
=====
--- ext/strscan/strscan.c    (revision 20570)
+++ ext/strscan/strscan.c    (working copy)
@@ -964,4 +964,90 @@ strscan_aref(VALUE self, VALUE idx)

/*
+ * call-seq: size
+ *
+ * Return the amount of subgroups in the most recent match.
+ * The full match counts as a subgroup.
+ *
+ *   s = StringScanner.new("Fri Dec 12 1975 14:39")
+ *   s.scan(/(\w+) (\w+) (\d+) /)      # -> "Fri Dec 12 "
+ *   s.size                            # -> 4
+ */
+static VALUE
+strscan_size(VALUE self)
+{
+  struct strscanner *p;
+
+  GET_SCANNER(self, p);
+  if (! MATCHED_P(p)) return Qnil;
+  return INT2FIX(p->regs.num_regs);
+}
+
+/*
+ * call-seq: captures
+ *
+ * Returns the subgroups in the most recent match (not including the full match).
+ * If nothing was priorly matched, it returns nil.
+ *
+ *   s = StringScanner.new("Fri Dec 12 1975 14:39")
+ *   s.scan(/(\w+) (\w+) (\d+) /)      # -> "Fri Dec 12 "
+ *   s.captures                        # -> ["Fri", "Dec", "12"]
+ *   s.scan(/(\w+) (\w+) (\d+) /)      # -> nil
+ *   s.captures                        # -> nil
+ */
+static VALUE
+strscan_captures(VALUE self)
+{
+  struct strscanner *p;
+  int i, num_regs;
+  VALUE new_ary;
+
+  GET_SCANNER(self, p);
+  if (! MATCHED_P(p)) return Qnil;
+
+  num_regs = p->regs.num_regs;
+  new_ary = rb_ary_new2(num_regs);
+
+  for (i = 1; i < num_regs; i++) {
+    VALUE str = extract_range(p, p->prev + p->regs.beg[i],
+                             p->prev + p->regs.end[i]);
+    rb_ary_push(new_ary, str);
+  }
+
+  return new_ary;
+}
+
+/*
+ * call-seq:
+ *   scanner.values_at( i1, i2, ... iN )  -> an_array
+ *
+ * Returns the subgroups in the most recent match at the given indices.
+ * If nothing was priorly matched, it returns nil.
+ *
+ *   s = StringScanner.new("Fri Dec 12 1975 14:39")
+ *   s.scan(/(\w+) (\w+) (\d+) /)      # -> "Fri Dec 12 "
+ *   s.values_at 0, -1, 5, 2          # -> ["Fri Dec 12 ", "12", nil, "Dec"]
+ *   s.scan(/(\w+) (\w+) (\d+) /)      # -> nil
+ *   s.captures                        # -> nil
+ */
+static VALUE
```

```

+strscan_values_at(int argc, VALUE *argv, VALUE self)
+{
+  struct strscanner *p;
+  long i;
+  VALUE new_ary;
+
+  GET_SCANNER(self, p);
+  if (! MATCHED_P(p)) return Qnil;
+
+  new_ary = rb_ary_new2(argc);
+  for (i = 0; i<argc; i++) {
+    rb_ary_push(new_ary, strscan_aref(self, argv[i]));
+  }
+
+  return new_ary;
+}
+
+/*
+ * Return the <i><b>pre</b>-match</i> (in the regular expression sense) of the last scan.
+ *
@@ -1312,4 +1398,7 @@ Init_strscan(void)
+  rb_define_method(StringScanner, "pre_match", strscan_pre_match, 0);
+  rb_define_method(StringScanner, "post_match", strscan_post_match, 0);
+  rb_define_method(StringScanner, "size", strscan_size, 0);
+  rb_define_method(StringScanner, "captures", strscan_captures, 0);
+  rb_define_method(StringScanner, "values_at", strscan_values_at, -1);
+
+  rb_define_method(StringScanner, "rest", strscan_rest, 0);

```

--
Nobu Nakada

#2 - 12/10/2008 05:01 AM - apeiros (Stefan Rusterholz)

Hi Nobu

Kind thanks for taking a look at my patch.

On 08.12.2008 at 18:27 Nobuyoshi Nakada wrote:
| Variable declaration is disallowed after executing statements.

Ah, good to know. I hope I will remember next time!
Can I setup extconfig.rb so that gcc will warn me if I did that?

| call-seq: for captures was `size`.

Whoops. "Too close" syndrome I guess :) Thanks for pointing that out.

#3 - 02/02/2009 11:57 AM - shyouhei (Shyouhei Urabe)

- Assignee set to nobu (Nobuyoshi Nakada)

#4 - 04/02/2010 08:08 AM - znz (Kazuhiro NISHIYAMA)

- Target version set to 2.0.0

#5 - 06/03/2010 10:45 AM - usa (Usaku NAKAMURA)

- Status changed from Open to Assigned

#6 - 07/27/2011 04:31 PM - drbrain (Eric Hodel)

- Category set to lib

#7 - 02/13/2012 08:54 PM - mame (Yusuke Endoh)

Nobu, do you think that this proposal can be accepted?
There is no maintainer for strscan; I leave this decision
to your discretion.

--
Yusuke Endoh mame@tsg.ne.jp

#8 - 10/27/2012 07:25 AM - nobu (Nobuyoshi Nakada)

Seems fine to me.

#9 - 11/24/2012 01:11 PM - mame (Yusuke Endoh)

- Target version changed from 2.0.0 to 2.6

#10 - 11/29/2017 07:57 AM - nobu (Nobuyoshi Nakada)

- Status changed from Assigned to Closed

Applied in changeset [trunk|r60929](#).

strscan.c: add MatchData-like methods

- ext/strscan/strscan.c: added size, captures and values_at to StringScanner, shorthands of accessing the matched data. based on the patch by apeiros (Stefan Rusterholz) at [ruby-core:20412]. [Feature [#836](#)]

Files

| | | | |
|--------------|---------|------------|-----------------------------|
| strscan.c | 38.2 KB | 12/08/2008 | apeiros (Stefan Rusterholz) |
| strscan.diff | 3.12 KB | 12/08/2008 | apeiros (Stefan Rusterholz) |