Reduce duplicated warnings for the change of Ruby 3 keyword arguments

11/01/2019 03:29 AM - mame (Yusuke Endoh)

Problem

Currently, the interpreter emits 200 lines of warnings against the following program.

```ruby
def foo(**opt); end
100.times { foo({kw:1}) }
```

$ ./miniruby -e 'def foo(**opt); end; 100.times { foo({kw:1}) }'
-e:1: warning: The last argument is used as the keyword parameter
-e:1: warning: for `foo' defined here
-e:1: warning: The last argument is used as the keyword parameter
-e:1: warning: for `foo' defined here
-e:1: warning: The last argument is used as the keyword parameter
-e:1: warning: for `foo' defined here
...

In theory, the warnings are not harmful because they don't stop or interfere the execution. But in practice, I'm afraid if they are annoying because they flush all console logs away. I think that the warning is not needed if the call is already warned.

Proposal

How about limiting the count of warnings to at most once for each pair of caller and callee?

I've created a pull request. It records all pairs of caller position and callee iseq when emitting a warning, and suppress the warning if the same pair of caller and callee is already warned.

What do you think?

Related issues:

Related to Ruby master - Feature #16345: Don't emit deprecation warnings by d...
Closed

Associated revisions

Revision 191ce534 - 11/29/2019 08:32 AM - mame (Yusuke Endoh)
Reduce duplicated warnings for the change of Ruby 3 keyword arguments

By this change, the following code prints only one warning.

```ruby
def foo(**opt); end
100.times { foo({kw:1}) }
```

A global variable st_table *caller_to_callees is a map from caller to a set of callee methods. It remembers that a warning is already printed for each pair of caller and callee.

[Feature #16289]

History

#1 - 11/01/2019 04:38 AM - marcandre (Marc-Andre Lafortune)
Definite +1 for me. I thought there was already an issue about this but looks I was mistaken.

#2 - 11/01/2019 07:03 AM - nobu (Nobuyoshi Nakada)
It leads a memory leak.
And different warnings won't be suppressed too?

#3 - 11/01/2019 01:26 PM - mame (Yusuke Endoh)

It leads a memory leak.

Theoretically, yes. Even if iseqs of caller or callee are free’d, the corresponding records are not free’d. But honestly I don't think it is worth elaborating the feature for some reasons:

(1) in practice, the case would be rare (to cause the leak, we need to repeatedly create iseq by using "eval" and trigger a warning in the iseq);
(2) the leak does not occur if a warning is not emitted (if any, it should be fixed); and
(3) this hack is only for Ruby 2.7.

Before implementing the patch, I discussed with ko1. At first I was going to avoid the leak by adding some fields to iseq, but he objected the approach and recommended the above design. Now I agree with him.

And different warnings won't be suppressed too?

It was not intended, but it would be also very rare to trigger different keyword warnings in one call, I guess.

#4 - 11/02/2019 08:00 AM - mame (Yusuke Endoh)

What I worry about is Jeremy's comment in the PR:

In terms of idea, whether we want this feature depends on how annoying we want the warnings to be. If they warn every call, they are very annoying and it will push people to fixing the issue. If they only warn once, they are not that annoying, and people may be less inclined to fix the issue.

It is really convincing. That being said, I'm afraid if it is too annoying.

I'd like to hear opinions and discuss it at the next meeting, so I've created this ticket.

#5 - 11/02/2019 12:56 PM - Eregon (Benoit Daloze)

What's the performance with this patch for a call site that would warn?

```
$ chruby ~2.7.0-preview2
$ ruby -rbenchmark/ips -e 'def m(**kw); end; h = {a: 1}; Benchmark.ips { |x| x.report { m(h) } }'
... many warnings ...
  88.603k (±15.8%) i/s - 434.163k in 5.025597s

$ ruby -rbenchmark/ips -e 'def m(**kw); end; h = {a: 1}; Benchmark.ips { |x| x.report { m(h) } }' 2>/dev/null
Warming up --------------------------------------
  46.040k i/100ms
Calculating -------------------------------------
  504.460k (± 7.0%) i/s - 2.532M in 5.050515s

$ chruby this-branch
$ ruby -rbenchmark/ips -e 'def m(**kw); end; h = {a: 1}; Benchmark.ips { |x| x.report { m(h) } }'
Warming up --------------------------------------
-e:1: warning: The last argument is used as the keyword parameter
-e:1: warning: for 'm' defined here
  372.010k i/100ms
Calculating -------------------------------------
  8.166M (± 2.6%) i/s - 40.921M in 5.015298s

$ chruby 2.6.2
$ ruby -rbenchmark/ips -e 'def m(**kw); end; h = {a: 1}; Benchmark.ips { |x| x.report { m(h) } }'
Warming up --------------------------------------
Calculating -------------------------------------
  322.637k i/100ms
  5.925M (± 8.8%) i/s - 29.683M in 5.055851s
```

#6 - 11/02/2019 03:47 PM - marcandre (Marc-Andre Lafortune)
If they only warn once, they are not that annoying, and people may be less inclined to fix the issue.

It's a warning about an upcoming incompatibility, once should be serious enough!

Maybe it should be reworded to emphasize seriousness, using the word "Deprecated" and "This will not be supported in Ruby 3.x":

```ruby
# Change:
warning: The last argument is used as the keyword parameter
warning: for `foo' defined here

# to something like this:
Deprecated: The last argument is used as the keyword parameter
Deprecated: for `foo' defined here. This will not be supported in Ruby 3.x
```

#7 - 11/04/2019 04:41 PM - Dan0042 (Daniel DeLorme)

If they only warn once, they are not that annoying, and people may be less inclined to fix the issue.

It's a warning about an upcoming incompatibility, once should be serious enough!

I think the issue is more that in a long-lived process you'll get only a few warnings in a large log, and it may be easy to miss unless you specifically grep for the warnings.

#8 - 11/05/2019 09:35 AM - shevegen (Robert A. Heiler)

Makes sense to me what mame wrote, so +1; marcandre's suggestion for a deprecation-specific notice makes sense as well and may be useful past ruby 3.0 (if ruby 4.x may have more incompatibilities than ruby 3.0; then there could be more distinct "Deprecated" or "Deprecation" messages).

Daniel wrote:

I think the issue is more that in a long-lived process you'll get only a few warnings in a large log, and it may be easy to miss unless you specifically grep for the warnings.

While your comment is perfectly reasonable and makes sense as well, I think this taps into the older discussion about customizing warnings issued by ruby, so that people can adjust ruby to different use cases (a bit like you can customize rubocop). But this may be a discussion past ruby 3.0 perhaps. IMO what mame wrote makes sense too since it may not be too overly useful to have your full screen overflow with the same warning-message spam that you already know is a problem. Of course if there was some way to customize it in a simple manner, ruby users could then adjust this to their general use case - but I don't want to dilute mame's suggestion here. Picking the "right" default behaviour covering the most likely or "best" use case of course makes a lot of sense to me and I think most ruby users who run into this situation are more likely to want to read cozy and helpful messages on the commandline, and using logs only secondary. :)

marcandre's suggestion has the advantage that the leading word "deprecated" can be colourized quite easily, e. g. people could "style" it to red on their computer; KDE Konsole supports R,G,B style too, I use that a lot myself for different colours (HTML colours such as "tomato" or "orangered", but I am going off-topic here a little bit, it is also not available to everyone, of course). But in general I think mame's description makes a lot of sense.

Note that while I understand jeremy's comment about annoying messages "helping" people change their code, not everyone thinks that spam is that helpful. See also past discussions about reversing the output of the errors so that some people have it easier to find the problem instantly, on the last part of the message, whereas other ruby users prefer the old style. In the long run I think it would be best to have some larger proposal for people being able to customize ruby warnings (and perhaps error output to, to some extent), but it should be simple to use too. But again, sorry for diluting here, +1 to what mame wrote. I don't think "nagging" should be of higher priority than the goal of making it simpler to read and understand what is going on - that is also in the spirit of the did-you-mean gem, if you ask me.

#9 - 11/10/2019 11:24 AM - Eregon (Benoit Daloze)

I updated the benchmark results above and measured on this branch.
It's ~100x faster in that micro benchmark.

If we don't want pretty bad performance when migrating to 2.7 without fixing kwargs warnings (which probably takes a long time for application with many dependencies), then I think we need this.

#10 - 11/10/2019 10:20 PM - sam.saffron (Sam Saffron)

Just to provide some context on the extent of the issue.

Running the spec suite for Discourse results in 2,698,774 rows being printed to STDOUT.

sort log | uniq -c | sort -bgr

https://gist.github.com/SamSaffron/7c29324a4a6a32a042d2c1c60e6ca8a2

It's ~100x faster in that micro benchmark.

If we don't want pretty bad performance when migrating to 2.7 without fixing kwargs warnings (which probably takes a long time for application with many dependencies), then I think we need this.

#10 - 11/10/2019 10:20 PM - sam.saffron (Sam Saffron)

Just to provide some context on the extent of the issue.

Running the spec suite for Discourse results in 2,698,774 rows being printed to STDERR.

sort log | uniq -c | sort -bgr

https://gist.github.com/SamSaffron/7c29324a4a6a32a042d2c1c60e6ca8a2
52  /home/sam/Source/discourse/app/models/reviewable.rb:339: warning: for `list_for' defined here
52  /home/sam/.rbenv/versions/trunk/lib/ruby/gems/2.7.0/gems/sassc-2.0.1/lib/sassc/script/value/string.rb:26: warning: deprecated Object== is called on TrueClass; it always returns nil
48  /home/sam/.rbenv/versions/trunk/lib/ruby/gems/2.7.0/gems/multi_json-1.13.1/lib/multi_json/options_cache.rb:12: warning: Capturing the given block using Proc.new is deprecated; use 'block' instead
46  /home/sam/.rbenv/versions/trunk/lib/ruby/gems/2.7.0/gems/activerecord-6.0.1/lib/active_record/inflector/methods.rb:174: warning: for `titleize_without_cache' defined here
44  /home/sam/.rbenv/versions/trunk/lib/ruby/gems/2.7.0/gems/sassc-2.0.1/lib/sassc/script/value/string.rb:26: warning: deprecated Object== is called on FalseClass; it always returns nil
36  /home/sam/.rbenv/versions/trunk/lib/ruby/gems/2.7.0/gems/actionpack-6.0.1/lib/action_dispatch/middleware/re/static.rb:110: warning: for `initialize' defined here
36  /home/sam/.rbenv/versions/trunk/lib/ruby/gems/2.7.0/gems/actionpack-6.0.1/lib/action_dispatch/middleware/re/stack.rb:37: warning: The last argument is used as the keyword parameter
34  /home/sam/.rbenv/versions/trunk/lib/ruby/gems/2.7.0/gems/excon-0.64.0/lib/excon.rb:178: warning: Capturing the given block using Proc.new is deprecated; use 'block' instead
32  /home/sam/Source/discourse/lib/seed_data/categories.rb:98: warning: for `create_category' defined here
32  /home/sam/Source/discourse/lib/seed_data/categories.rb:15: warning: The last argument is used as the keyword parameter
26  /home/sam/Source/discourse/app/controllers/reviewables_controller.rb:331: warning: The last argument is used as the keyword parameter
26  /home/sam/Source/discourse/app/controllers/reviewables_controller.rb:331: warning: The last argument is used as the keyword parameter
22  /home/sam/.rbenv/versions/trunk/lib/ruby/gems/2.7.0/gems/onebox-1.9.22/lib/onebox/helpers.rb:222: warning: URI.escape is obsolete
72: warning: deprecated Object== is called on Class; it always returns nil
18  /home/sam/Source/discourse/lib/seed_data/topics.rb:16: warning: The last argument is used as the keyword parameter
18  /home/sam/Source/discourse/lib/seed_data/topics.rb:16: warning: The last argument is used as the keyword parameter
16  /home/sam/Source/discourse/lib/seed_data/categories.rb:24: warning: The last argument is used as the keyword parameter
16  /home/sam/Source/discourse/lib/seed_data/categories.rb:148: warning: for `update_category' defined here
14  /home/sam/Source/discourse/lib/compression/strategy.rb:35: warning: for `strip_directory' defined here
14  /home/sam/Source/discourse/lib/compression/engine.rb:25: warning: The last argument is used as the keyword parameter
14  /home/sam/Source/discourse/lib/compression/engine.rb:25: warning: The last argument is used as the keyword parameter
14  /home/sam/Source/discourse/lib/compression/engine.rb:25: warning: The last argument is used as the keyword parameter
12  /home/sam/.rbenv/versions/trunk/lib/ruby/gems/2.7.0/gems/activerecord-6.0.1/lib/activerecord/connection_adapters/postgresql_adapter.rb:620: warning: given argument is nil; this will raise a TypeError in the next release
10  /home/sam/Source/discourse/app/models/embeddable_host.rb:37: warning: URI.unescape is obsolete
8: warning: deprecated Object== is called on Float; it always returns nil
6: warning: deprecated Object== is called on TrueClass; it always returns nil
4: warning: for `initialize' defined here
2: warning: for `transaction' defined here
1: warning: for `initialize' defined here
1: warning: for `transaction' defined here
1: warning: for `initialize' defined here
1: warning: for `transaction' defined here
1: warning: for `initialize' defined here
1: warning: for `transaction' defined here
1: warning: for `initialize' defined here
1: warning: for `transaction' defined here

Honestly I am not sure if the per call-site thing is needed, I would almost just have a default of count everything and just stop once you reach a number.

Pick a number 1000, 5000 ... something ... once you reach that ... you print to STDERR "Too many warnings regarding keyword args reached, you
will no longer be alerted". You could then start small at like 50 for 2.7 release and crank it up to 500 at 2.7.1 if we feel people are not being bugged enough.

Technically even if the limit was 1 line globally ... that is enough to get an entire app fixed.

At the current levels I just see people using RUBYOPT=-W0 cause often they will require complex fixes to gems beyond an app owners controls (especially now that delegation is a nightmare)

#11 - 11/11/2019 01:00 AM - sam.saffron (Sam Saffron)
Note... some of this fiddling with keyword arguments is showing up as a slowdown in rubybench

https://rubybench.org/ruby/ruby/commits?result_type=app_answer&display_count=2000

#12 - 11/13/2019 07:25 AM - duerst (Martin Dürst)
- Related to Feature #16345: Don't emit deprecation warnings by default. added

#13 - 11/29/2019 08:12 AM - mame (Yusuke Endoh)
Matz accepted this with #16345, I'll commit it soon.

#14 - 11/29/2019 08:32 AM - mame (Yusuke Endoh)
- Status changed from Open to Closed

Applied in changeset g3c191c534ce42c91571f47c85be913826b1c7.

Reduce duplicated warnings for the change of Ruby 3 keyword arguments

By this change, the following code prints only one warning.

def foo(**opt); end
100.times { foo(kw:1) }

A global variable st_table *caller_to_callees is a map from caller to a set of callees methods. It remembers that a warning is already printed for each pair of caller and callee.

[Feature #16289]

#15 - 12/23/2019 07:37 AM - marcandre (Marc-Andre Lafortune)
I'm glad this was merged in, but disappointed the error message still doesn't mention that this is a deprecation warning.

Wouldn't it be cleared with https://github.com/ruby/ruby/pull/2776 ?

#16 - 12/23/2019 11:49 AM - mame (Yusuke Endoh)
- Assignee set to mame (Yusuke Endoh)
- Status changed from Closed to Assigned

marcandre (Marc-Andre Lafortune), thank you for creating a pull request.

Your proposal:

old: The last argument is used as keyword parameters; maybe ** should be added to the call
new Using the last argument as keyword parameters is deprecated; maybe ** should be added to the call

old: The keyword argument is passed as the last hash parameter
new: Passing the keyword argument as the last hash parameter is deprecated

old: The last argument is split into positional and keyword parameters
new: Splitting the last argument into positional and keyword parameters is deprecated

I think they are better. I'm okay to merge your patch, if we can merge it before 2.7.0 release. (But I'm against changing them after the release, i.e., 2.7.1 or later.)

matz (Yukihiro Matsumoto) jeremyevans@ (Jeremy Evans) Eregon (Benoit Daloze) What do you think?

naruse (Yui NARUSE) If matz accepted, can we merge it?
#17 - 12/23/2019 12:51 PM - Eregon (Benoit Daloze)
Sounds clearer to me with marcandre (Marc-Andre Lafortune)'s changes.

#18 - 12/23/2019 05:47 PM - jeremyevans0 (Jeremy Evans)
I'm fine with marcandre (Marc-Andre Lafortune)'s changes.

#19 - 12/23/2019 09:49 PM - marcandre (Marc-Andre Lafortune)
- Status changed from Assigned to Closed

Great, thanks.
Pushed.

#20 - 12/24/2019 07:06 AM - mame (Yusuke Endoh)
marcandre (Marc-Andre Lafortune)
You have to wait for matz and naruse's approvals! I'll ask them.

#21 - 12/24/2019 07:10 AM - matz (Yukihiro Matsumoto)
Approved.
Matz.

#22 - 12/24/2019 07:23 AM - mame (Yusuke Endoh)
naruse (Yui NARUSE) also approved. Thank you for quick replies.

#23 - 12/24/2019 07:34 AM - mame (Yusuke Endoh)
I've updated my pull request to update the article of Ruby 3's keyword argument separation.
https://github.com/ruby/www.ruby-lang.org/pull/2324