Сложные SQL запросы в Ruby On Rails

Манылов Павел aka [r-k] SFD 2012 / Екатеринбург

Сложные запросы?

- Не описаны в туториалах и документации по Rails
- Чаще всего затрагивают больше одной таблицы
- Вызывают затруднения при написании

Короткий обзор

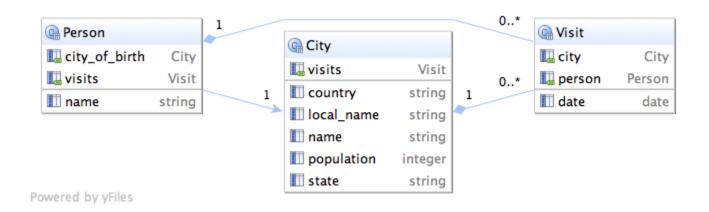
- Rails MVC фреймворк
- ActiveRecord ORM библиотека
- ARel менеджер синтаксических деревьев для составления sql-запросов

Rails использует ActiveRecord использует ARel

Требования к коду

- Красивость
- Читабельность
- Без вставок SQL-кода
- Гибкость, DRY
- Rails Way

Используем следующую структуры БД



Необходимо получить население по областям (штатам) в виде массива объектов со следующими данными:

```
[{
    :country=>"Страна",
    :state=>"Область",
    :population=>20000
},
...]
```

```
City.group(:country,:state) \
    .sum(:population)
```

```
City.group(:country,:state) \
    .sum(:population)
# => { ["Страна", "Область"] => 10000}
```

```
City.group(:country,:state) \
    .sum(:population) \
    .map{|k,v|}
            :country=>k[0],
            :state=>k[1],
            :population=>v
     } # => Array<Hash>
```

```
City.group(:country,:state) \
 .sum(:population) \
 <u>-map{|k,v|</u>
             :country=>k[0]_{r}
            :state=>k[1],
            :population=>v
      <del>} # => Array<Hash></del>
```

```
at = City.arel_table
City.find_by_sql(
   at.project(
      at[:country],
      at[:state], at[:population].sum.as
      ("population")
   ).group(at[:country],at[:state])
)
```

```
at = City.arel_table
City.find_by_sql(
   at.project(
      at[:country],
      at[:state], at[:population].sum.as
      ("population")
   ).group(at[:country],at[:state])
```

```
at = City.arel_table
City.find_by_sql(
   at.project( # Arel::SelectManager.new(...).from(...).project(...)
   at[:country],
   at[:state], at[:population].sum.as
    ("population")
).group(at[:country],at[:state])
)
```

```
at = City.arel_table
City.find_by_sql(
   at.project(
      at[:country],
      at[:state], at[:population].sum.as
      ("population")
   ).group(at[:country],at[:state])
)
```

```
at = City.arel_table
City.find_by_sql(
   at.project(
      at[:country],
      at[:state], at[:population].sum.as
         ("population")
   ).group(at[:country],at[:state])
)
```

```
at = City.arel_table
City.find_by_sql(
   at.project(
      at[:country],
      at[:state], at[:population].sum.as
      ("population")
).group(at[:country],at[:state])
)
```

ActiveRecord + ARel:

ActiveRecord + ARel:

Необходимо получить список городов, которые посещали люди, посещавшие так же город X, в виде массива объектов модели City:

```
[<City>, ...]
```

```
City.find_by_sql("
SELECT * FROM `cities`
WHERE `cities`.`id` IN (
    SELECT `v1`.`city_id` FROM `visits` as `v1`
    WHERE `v1`.`person_id` IN(
        SELECT `v2`.`person_id` FROM `visits` as `v2`
        WHERE `city_id` = #{city_id}
    )
)
")
```

```
City.find_by_sql("
SELECT * FROM `cities`
WHERE `cities`.`id` IN (
    SELECT `v1`.`city_id` FROM `visits` as `v1`
    WHERE `v1`.`person_id` IN(
        SELECT `v2`.`person_id` FROM `visits` as `v2`
        WHERE `city_id` = #{city_id}
    )
)
")
```

```
City.find_by_sql("
SELECT * FROM `cities`
WHERE `cities`.`id` IN (
    SELECT `v1`.`city_id` FROM `visits` as `v1`
    WHERE `v1`.`person_id` IN(
        SELECT `v2`.`person_id` FROM `visits` as `v2`
        WHERE `city_id` = #{city_id}
    )
)
")
```

```
City.find by_sql(
   at.project(Arel.star).where(
      at[:id].in(
          v1.project(v1 alias[:city id]).where(
             v1 alias[:person id].in(
                 v2.project(v2_alias[:person_id]).where(
                    v2 alias[:city_id].eq(city_id)
```

```
at.project("*")
                              at.project(Arel::SqlLiteral.new("*"))
                              # .to sql => "SELECT * FROM `cities`"
ARel:
                              at.project(at["*"])
City.find_by sql(
                              at.project(at[Arel.star])
                              # .to sql => "SELECT `cities`.* FROM `cities`"
    at.project (Arel.star) .where (
        at[:id].in(
            v1.project(v1 alias[:city id]).where(
                v1 alias[:person id].in(
                    v2.project(v2 alias[:person id]).where(
                        v2 alias[:city_id].eq(city_id)
```

ARel:

```
City.find by sql(
   at.project(Arel.star).where(
       at[:id].in(
           v1.project(v1 alias[:city id]).where(
               v1 alias[:person id].in(
                   v2.project(v2 alias[:person id]).where(
                      v2 alias[:city id].eq(city_id)
                  Arel::Nodes::Equality.new(v2 alias[:city id], cities[1].id)
                  # .to sql => "`v2`.`city id` = X"
```

ARel: Arel::Nodes::In.new(at[:id], v1....) City.find by sql(# .to sql => "`cities`.`id` IN (SELECT ...)" at.project(Arel.star).where(at[:id]. **in** (v1.project(v1 alias[:city id]).where(v1 alias[:person id] **←in**(v2.project(v2 alias[:person id]).where(v2 alias[:city_id].eq(city_id)

ARel:

НЕ ОЧЕНЬ

Я думал, намного будет... Намного лучше будет это все

Сколько раз сюда ходи-и-и-л — было намного лучше, но на этот раз как-то не удало-о-ось.



ActiveRecord only:

ActiveRecord only:

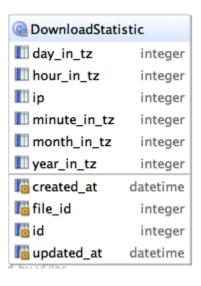
ActiveRecord only:

ActiveRecord only:

SO GOOD!



Достаточно часто используемый пример: статистика скачиваний файлов.



Необходимо получить *статистику* уникальных загрузок файлов с группировкой по дням в виде:

```
DownloadStatistic.select([:file id, :year in tz,
                           :month in tz, :day in tz,
                          Arel.star.count.as("hosts count")]) \
                 .from( DownloadStatistic.select(
                           [:file id,:year in tz, :month in tz, :day in tz]
                        ).group(:file id, :ip,
                                :year in tz,:month in tz, :day in tz) \
                           arel.as("t1")
                 ).group(:file id, :year in tz, :month_in_tz, :day_in_tz)
 <Arel::SelectManager>
```

JOIN-ы таблиц

Возьмём в качестве примера интернетмагазин. В данный момент нас интересуют следующие таблицы и связь между ними:



JOIN-ы таблиц

Необходимо получить массив объектов моделей Client с дополнительными полями, отражающими общее количество покупок и сумму потраченных в магазине средств в виде:

Иногда необходимо динамически добавлять к запросу поля выборки:

```
class ActiveRecord::Relation
  def add_select(selects)
    t=self.scoped
    selects=[selects] unless selects.kind_of?(Array)
    t.select_values = [self.arel_table[Arel.star]] if
t.select_values.blank?
    t.select_values+=selects
    t
  end
end
```

Иногда необходимо динамически добавлять к запросу поля выборки:

```
class ActiveRecord::Relation
  def add_select(selects)
    t=self.scoped
    selects=[selects] unless selects.kind_of?(Array)
    t.select_values = [self.arel_table[Arel.star]] if
t.select_values.blank?
    t.select_values+=selects
    t
  end
end
```

Примеры использования:

```
p = Purchase.select(Purchase.arel_table[:sum])
# .to_sql => SELECT `purchases`.`sum` FROM `purchases`

x = p.add_select(Purchase.arel_table[:client_id])
# . to_sql => SELECT `purchases`.`sum`, `purchases`.
   `client_id` FROM `purchases`
```

Примеры использования:

```
p = Purchase.select(Purchase.arel_table[:sum])
# .to_sql => SELECT `purchases`.`sum` FROM `purchases`

x = p.add_select(Purchase.arel_table[:client_id])
# . to_sql => SELECT `purchases`.`sum`, `purchases`.
`client id` FROM `purchases`.
```

Примеры использования:

```
p = Purchase.select(Purchase.arel_table[:sum])
# .to_sql => SELECT `purchases`.`sum` FROM `purchases`

x = p.add_select(Purchase.arel_table[:client_id])
# . to_sql => SELECT `purchases`.`sum`, `purchases`.
`client id` FROM `purchases`
```

```
p at = Purchase.arel table
c at = Client.arel table
purchases join = Arel::OuterJoin.new(
   p at,
   Arel::Nodes::On.new(
     c at[:id].eq(p at[:client id])
#.to sql => "LEFT OUTER JOIN `purchases` ON `clients`.`id`
= `purchases`.`client id`"
```

```
p at = Purchase.arel table
c at = Client.arel table
purchases join = Arel::OuterJoin.new(
   p at,
   Arel::Nodes::On.new(
     c at[:id].eq(p at[:client id])
#.to sql => "LEFT OUTER JOIN `purchases` ON `clients`.`id`
= `purchases`.`client id`"
```

```
p at = Purchase.arel table
c at = Client.arel table
purchases join = Arel::OuterJoin.new(
   p_at,
   Arel::Nodes::On.new(
     c at[:id].eq(p at[:client id])
#.to sql => "LEFT OUTER JOIN `purchases` ON `clients`.`id`
= `purchases`.`client id`"
```

```
p at = Purchase.arel table
c at = Client.arel table
purchases join = Arel::OuterJoin.new(
   p at,
   Arel::Nodes::On.new(
     c at[:id].eq(p at[:client id])
#.to sql => "LEFT OUTER JOIN `purchases` ON `clients`.`id`
= `purchases`.`client_id`"
```

```
#p at, c at, purchases join определили ранее
Client.joins(purchases join) \
       .add select(
         p at[:id].count \
         .as("purchases count")
       ).group(p at[:client id])
# .to sql => "SELECT `clients`.*, COUNT(`purchases`.`id`)
AS purchases count FROM `clients` LEFT OUTER JOIN
`purchases` ON `clients`.`id` = `purchases`.`client id`
GROUP BY `purchases`.`client id`"
```

```
#p at, c at, purchases join определили ранее
Client.joins(purchases join) \
       .add select(
         p at[:id].count \
         .as("purchases count")
       ).group(p at[:client id])
# .to sql => "SELECT `clients`.*, COUNT(`purchases`.`id`)
AS purchases count FROM `clients` LEFT OUTER JOIN
`purchases` ON `clients`.`id` = `purchases`.`client id`
GROUP BY `purchases`.`client id`"
```

```
#p at, c at, purchases join определили ранее
Client.joins(purchases join) \
       .add select(
         p at[:id].count \
         .as("purchases count")
       ).group(p at[:client id])
# .to sql => "SELECT `clients`.*, COUNT(`purchases`.`id`)
AS purchases count FROM `clients` LEFT OUTER JOIN
`purchases` ON `clients`.`id` = `purchases`.`client id`
GROUP BY `purchases`.`client id`"
```

```
#p at, c at, purchases join определили ранее
Client.joins(purchases join) \
       .add select(
         p at[:id].count \
         .as("purchases count")
       ).group(p at[:client id])
# .to sql => "SELECT `clients`.*, COUNT(`purchases`.`id`)
AS purchases count FROM `clients` LEFT OUTER JOIN
`purchases` ON `clients`.`id` = `purchases`.`client id`
GROUP BY `purchases`.`client id`"
```

Аналогично для spent_money:

Аналогично для spent_money:

joins_purchase в виде named scope с предотвращением повторных join-ов:

```
class Client < ActiveRecord::Base
  scope :joins purchase, lambda{
    unless defined?(@joins purchase)
      p = Purchase.arel table
      @joins purchase = joins(Arel::OuterJoin.new(p, Arel::
Nodes::On.new(arel table[:id].eq(p[:client_id])))).group(p
[:client id])
    end
    @joins purchase
```

```
class Client < ActiveRecord::Base
 scope :with purchases count, lambda{
    joins purchase.add select(
      Purchase.arel table[:id].count.as("purchases count")
 scope :with spent money, lambda{
    joins purchase.add select(
      Purchase.arel table[:sum].sum.as("spent money")
```

```
class Client < ActiveRecord::Base
 scope :with purchases count, lambda{
    joins purchase.add select(
      Purchase.arel table[:id].count.as("purchases count")
 scope :with spent money, lambda{
    joins purchase.add select(
      Purchase.arel table[:sum].sum.as("spent money")
```

```
class Client < ActiveRecord::Base
 scope :with purchases count, lambda{
    joins purchase.add select(
      Purchase.arel table[:id].count.as("purchases count")
 scope :with spent money, lambda{
    joins purchase.add select(
      Purchase.arel table[:sum].sum.as("spent money")
```

JOIN-ы таблиц: Scopes

```
class Client < ActiveRecord::Base</pre>
  scope :with purchases information, lambda{
    with purchases count.with spent money
end
Client.with purchases information
# .to sql => "
#SELECT `clients`.*, COUNT(`purchases`.`id`) AS purchases count,
        SUM(`purchases`.`sum`) AS spent money
#FROM `clients`
#LEFT OUTER JOIN `purchases` ON `clients`.`id` = `purchases`.`client id`
#GROUP BY `purchases`.`client id`
# "
```

JOIN-ы таблиц: Проблема

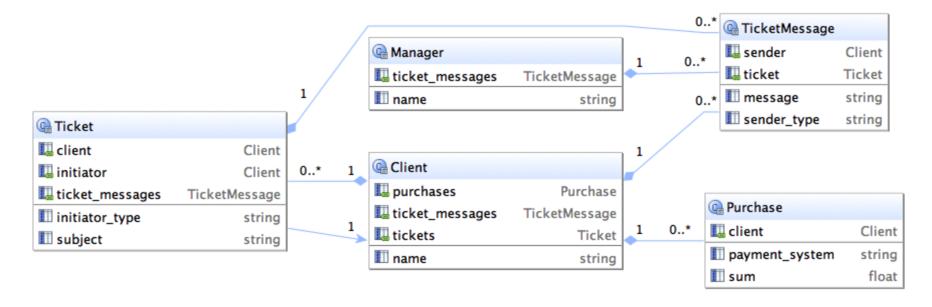
JOIN-ы таблиц: Проблема

Решения:

- Переписать со вложенными запросами
- Считать количество не COUNT(*) запросом, а на стороне Rails: Client.with_purchases_information. all.count

Polymorphic Associations

Используем тот же самый магазин, добавим к нему систему тикетов (общение менеджеров с клиентами)



Polymorphic Associations

```
class Ticket < ActiveRecord::Base</pre>
 belongs to :client
  belongs to :initiator, :polymorphic => true
  has many :ticket messages
end
class TicketMessage < ActiveRecord::Base</pre>
 belongs to :sender, :polymorphic => true
 belongs to :ticket
end
class Manager < ActiveRecord::Base</pre>
  has many :ticket messages, :as=>:sender
end
```

Polymorphic Associations

Необходимо получить массив сообщений для данного тикета с указанием имени отправителя в виде:

```
c at = Client.arel table
tm at = TicketMessage.arel table
sender_client_join = Arel::Nodes::OuterJoin.new(
  c at,
 Arel::Nodes::On.new(
    c at[:id].eq(tm at[:sender id]). \
    and(tm at[:sender type].eq("Client"))
```

```
c_at = Client.arel table
tm at = TicketMessage.arel table
sender client join = Arel::Nodes::OuterJoin.new(
  c at,
 Arel::Nodes::On.new(
    c at[:id].eq(tm at[:sender id]). \
    and(tm at[:sender type].eq("Client"))
```

```
m at = Manager.arel table
tm at = TicketMessage.arel table
sender manager join = Arel::Nodes::OuterJoin.new(
  m at,
  Arel::Nodes::On.new(
    m at[:id].eq(tm at[:sender id]). \
    and(tm at[:sender type].eq("Manager"))
```

```
m at = Manager.arel table
tm at = TicketMessage.arel table
sender manager join = Arel::Nodes::OuterJoin.new(
  m at,
  Arel::Nodes::On.new(
    m at[:id].eq(tm at[:sender id]). \
    and(tm at[:sender type].eq("Manager"))
```

Polymorphic Associations: Запрос

Итоговый запрос:

```
ticket = Ticket.find(some_id)
ticket.ticket_messages \
    .joins(sender_client_join).joins(sender_manager_join)
\
    .add_select(
        Arel::Nodes::NamedFunction.new("IFNULL",
        [
            c_at[:name],
            m_at[:name]
        ]).as("sender_name")
)
```

Polymorphic Associations: Запрос

Итоговый запрос:

```
ticket = Ticket.find(some id)
ticket.ticket messages \
       .joins(sender client join).joins(sender manager join)
       .add select(
        Arel::Nodes::NamedFunction.new("IFNULL",
           c at[:name],
           m at[:name]
         ]).as("sender name")
          # .to sql => "IFNULL(`clients`.`name`, `managers`.`name`) AS sender name"
```

Ну и хватит с запросами



Полезные ruby gems

- Squeel sql запросы в ruby стиле https://github.com/ernie/squeel
- Advanjo сложные join-ы в одну строку https://github.com/rap-kasta/advanjo.git
- Chooseme подбор индекса по используемым столбцам https://github.com/rap-kasta/chooseme.git

ВОПРОСЫ?

Контакты

Манылов Павел Сергеевич

email: rapkasta@gmail.com

jabber: rapkasta@jabber.ru

habrahabr: rapkasta

twitter: rap_kasta