



Compact Normal Form 已经可用



Felis sap...

函数式编程、编程语言、编程 话题的优秀回答者

已关注

开源哥、圆角骑士魔理沙、祖与占、草莓大福、hhhhhhhhh 等 29 人赞了该文章

Compact Normal Form 是即将到来的 ghc 8.2 一大特性。简而言之，对于满足特定条件的 heap object，rts 能够将其完全求值后放到一系列 region 中。这样做的好处包括：

- 可以直接通过对 region 进行 memcpy，完成对一系列 heap object 的序列化和反序列化。与传统的 binary serialization 框架如 binary、cereal 等相比，能够省去大量的解析工作，提高效率。
- 对于 long-lived 的 large heap object，即使在一轮 major gc 中也不需要再次遍历。
- 不需要 NFData 实例，也可以将 heap object 完全求值。NFData 提供的 deepseq 每次调用必定会遍历整个 object，而 CNF 的 compact 不用担心重复遍历。一般的 binary serialization 框架需要提供 get/put 的实现（虽然可以用 Generic 或者 TH 自动生成），而 CNF 连这步也省去了。

这项工作主要是 Edward Z. Yang 做的，详细可以参考他的论文 [Efficient Communication and Collection with Compact Normal Forms](#)。这里只提一下用法好了：

CNF 的接口分别在 ghc-compact 和 compact 两个包中提供，前者为 ghc 内置。ghc-compact 的 `GHC.Compact` 模块提供了接口，用于普通的 Haskell value 与其 compact 版本之间的转换。`compact :: a -> Compact a` 没有任何 constraint（嗯，没有 NFData），但是处理到带函数或者可变引用的值时会抛异常。由于 Haskell 的 non-strict 语义，带有 sharing 甚至 cycle 的值并不



至于对 region 进行序列化/反序列化，比较简单的做法是直接 Edward 的 compact 库提供的 `Data.Compact.Serialize`，可以将一个 region 从文件或者 Handle 进行读写。他的 compact 实现中用到了一些 magic number 和 Typeable 类提供的运行时类型信息做了一些动态的 sanity check。当然，可以用 ghc-compact 中的接口自己另外实现一个 serializer，比如我写的例子：gist.github.com/TerrorJ...，不需要操作文件，可以直接写到 in-memory buffer 中，更加方便一些。需要注意，serialize/deserialize 时，相应的 Compact 值一定要标注正确的类型，不然分分钟 Access violation。

自不用说，CNF 高性能的代价是 non-safety 和 non-portability。如果是从 untrusted source 获取 binary data，那还是用传统的框架更加保险，序列化出来的 region 也不可以移植到 ABI 不兼容的机器上。（package database 里面每个 package 版本号后面跟的一串 hash 代表一个 ABI）

等 8.2 正式版发布，我找个时间写个详细的 benchmark 对比一下各种 workload 下 CNF 和传统的序列化框架好了。

发布于 2017-05-19

「真诚赞赏，手留余香」

赞赏

还没有人赞赏，快来当第一个赞赏的人吧！

函数式编程

Haskell

GHC (编程套件)

▲ 赞同 29 ▼

● 3 条评论

➤ 分享

★ 收藏

...

文章被以下专栏收录



不动点高校现充部

一切与编程语言理论、函数式编程相关的杂谈。

已关注

推荐阅读





Notes on Haskell Debugging

elis... 发表于不动点高校...



幻想中的Haskell - Compiling Combinator

圆角骑士魔... 发表于雾雨魔法店



剖析Haskell应用架构

祖与占



函数

题叶

3 条评论

⇌ 切换为时间排序

写下你的评论...



李约瀚

1 年前

所以 8.2何时出。。

👍 2

以上为精选评论 ?



吴育昕

1 年前

orz edward yang...

👍 赞



祖与占

1 年前

能在这里[haskell-perf/serialization](#)搞?

👍 赞

