# Ravaで見る Rava Rubyで作ったJava VM Java仮想マシンのしくみ

東京農工大学大学院工学研究科 ささだ こういち Kohichi SASADA sasada@namikilab.tuat.ac.jp

## はじめに

本稿ではJava 仮想マシン(以下JVM)のしくみについて紹介します. ただ紹介するだけではつまらないので,プログラミング言語Ruby [1]で実装したJVMであるRava [2]の解説を交えながら進めていきます.本稿は, Java プログラムが実際にどうやってJVM上で動いているのか疑問に思っている方を対象にしているため,ある程度Java プログラムを書くことができるということを前提としています<sup>注1</sup>.

なお、本稿はJVMの詳細な仕様や実装までは述べません。そのため、自分でJVMを作ってみようという方は、必ず『Java仮想マシン仕様』[3]を参照してください(英語では全文オンラインで読めます[4]).
JVMを作ろうとまでは思わない方も、JVMの仕様を勉強することでJavaについての理解が深まると思うの

でおすすめです。本稿も基本的にこの仕様に基づいて 説明します。また,村山敏清さんが以前本誌で連載 していたJVMの解説記事[6]を含めた,組み込みJava の情報をWeb上で公開[6]されていますので,そちら もぜひ参考にしてください。

Java プログラムはどう動く?

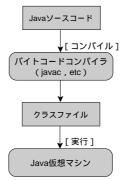
ご存知のとおり, Java プログラムを実行するためには,だいたい図1のような実行手順を踏みます.

Java で記述されたソースプログラムは , Java バイトコードコンパイラ (javac など) でコンパイルすることにより , クラスファイルへと変換されます . そして , クラスファイルのとおりにJVM は動作します . JVM はクラスファイルのみを相手にするため , Java 以外の言語で記述されたプログラムでも , それを適切にバイトコードへ変換するコンパイラさえあれば , JVM 上で実行することが可能です .

## 嗲 参考文献

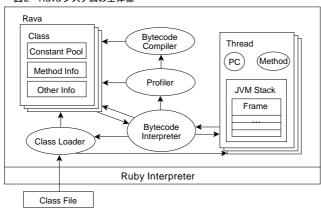
- [1] オブジェクト指向言語Ruby: http://www.ruby-lang.org/ja/
- [2] Rava / JavaVM on Ruby JPサポート用:http://www.namikilab.tuat.ac.jp/ sasada/prog/rava\_jp.html
- [3]『Java 仮想マシン仕様第2版』 / ティム・リンドホルム,フランク・イェリン著 / 村上雅章訳 / ピアソン・エデュケーション 2001
- [4] The Java Virtual Machine Specification: http://java.sun. com/docs/books/vmspec/
- [5]「Java 仮想マシン入門」/村山敏清著/『JAVA PRESS』 Vol.14~23連載/技術評論社
- [6]組込みJava情報源: http://www.netgene.co.jp/java/index.html

図1 Java プログラムが動くまで



注1 偉そうなこと書いてますが,きっと私より皆さんのほうがJava得意だと思います.

図2 Ravaシステムの全体像



そもそもなぜJVMが必要なのか、というのは使い古された説明で恐縮ですが、仮想マシンを実際の計算機システム上に被せることでハードウェアアーキテクチャやオペレーティングシステムなどから独立をはかる、つまり、どこでも一緒のプログラムを動かしたいという目的によります。また、JVM自体にセキュリティ機能をもたせることで実行の堅牢性を保証することができます。計算機システムにおいて、仮想マシンのような抽象層を1枚被せるとシステムは10倍遅くなる、というような話もあるそうですが、そのようなデメリットを差し引いても互換性や堅牢性が欲しかったわけです。もちろんJVMの実装技術が発達すれば、プラットフォーム依存のプログラムの実行速度とのギャップは縮まる(もしくは逆転する?)はずです。

Ravaって何?

Ravaとは,すべてRubyで記述した,Ruby100%, まじりっ気なしのJava 仮想マシンです<sup>注2</sup>.

従来JVM はCやC++で実装されてきました.筆者の知る限りでは,ほかにはJavaでの実装[7]しかありません.その最大の理由は実行速度です.前述したとおり,JVM という抽象層があるためシステムの処理速度が遅くなってしまうので,できるだけ実行速度を速くするようなプログラミング言語が選ばれるわけです.Ruby はお世辞にも処理速度に優れた実行環境があるとは言えません.

では,なぜRubyでRavaを作ったのか? その動機

は,誰もやってなかったからです.作るのが楽しかったからです.結果としては,処理速度としてはやはり遅いものになりましたが,JVMの理解は深まったと思いますし,この記事を書かせてもらうこともできました.

Rava のシステムの全体像としては,図2のようになります.

Rava を実行してみる

何はともあれ, Rava を実際に動かしてみましょう. Rava のインストール, セッティングについては本稿サポート用ページ[2]をご覧ください.

セットアップが終了したら,UNIX系OSならばシェル,Windowsならばコマンドプロンプト上でruby rava.rb Counterと入力してください.次のような出力が得られます.

\$ ruby rava.rb Counter

init Counter

init Counter

init Counter2

1

extended add

extended add

5

☞ 参考文献

[7] Jalapeno project: http://www.research.ibm.com/jalapeno/

注2 まだまだ仕様を満たしてない部分もあるのですが.

サンプルに利用したクラスCounter およびCounter2 を記述したソースをリスト1に示します. サンプルは 見てのとおりとても簡単なもので、クラス名にもやる 気が感じられません.しかし, Rava がこのサンプルく らいは実行可能であることを示しました.これくらい できなくてJVMを名乗るな,という気もします.

なお,プログラム中のrj.out()メソッドは,Rava 用のコンソール出力メソッドだと思ってください. ち

## リスト1 Counter.java

```
// Base Counter Class public class Counter{
     private int counter;
    public static void main(String[] args){
    Counter cnt = new Counter(); //
                                                  // オブジェクト生成
          Counter cnt2 = new Counter2();
         rj.out(cnt.add(1));
cnt2.add(2);
                                        // メソッド起動
// 派生クラスメソッド起動
         cnt2.add(3);
rj.out(cnt2.get());
                                        // 出力
    Counter(){
                                        // コンストラクタ
           counter = 0:
         rj.out("init Counter");
     public int add(int a){
                                        // ローカル変数アクセス・計算
// フィールドアクセス・計算
          for(int i=0;i<a;i++){
              counter++;
         return counter;
    public int get(){
     return counter;
// Extended Counter Class
public class Counter2 extends Counter{
Counter2(){ //派
                                        // 派生クラスコンストラクタ
         rj.out("init Counter2");
    public int add(int a){
    rj.out("extended add");
          return super.add(a);
                                        // 基底クラスメソッド呼び出し
```

なみに, Java 2 SDK (以下JDK) 付属のJava環境で java Counterのように実行しても,同様の出力を 得ることができます.

## クラスファイル

JVM がどのように動くのかを記述してあるのがクラ スファイルでした.ここでは,クラスファイルにはど んな情報が詰まっているかを概観します.

クラスファイルの中身を覗く

とりあえずクラスファイルをエディタで開いてみる と,よくわからないデータの並びが表示されます.こ れはクラスファイルがバイナリデータであるためです. そこで, JDK に付属しているクラスファイルのディス アセンブラである javap コマンドを利用します.こ のコマンドはクラスファイルを人間が読める形にして

では, Counter クラスのクラスファイルを覗いてみ ましょう.javap -c Counterと実行してみます

javap コマンドによって, ①Counter クラスに存在 するメソッドやフィールド,❷各メソッドの動作,な どがわかります.

### ☞ 参考文献

[8]『オブジェクト指向スクリプト言語 Ruby』/まつもとゆきひ ろ,石塚圭樹著/アスキー 1999

## プログラミング言語 Ruby

プログラミング言語 Ruby は,まつもとゆきひろさんに よってデザインされたオブジェクト指向なスクリプト言語で す.

Ruby は本格的なオブジェクト指向言語であり、たとえ ばクラスやその派生などの概念は本誌読者の方には馴染み 深いものであると思います. それに加え, Mix-in や特異メ ソッドなどをサポートしています. そして, Ruby はすべて がオブジェクトです.

インタプリタであるため、コンパイルなどの作業を行わず にRuby スクリプトを実行することが可能です.また,テ キスト処理の能力にも優れており, Perl と同じくらい強力

です. さらにシンプルな文法と, 例外処理やイテレータな どの機構によって、よりわかりやすいプログラミングを可能 にしています. その上, ガーベッジコレクタやスレッドなど を備えながら,移植性が高くさまざまなプラットフォームに 移植されています.

Ruby の文法を学習するためには俗にいうRuby 本[8] (オススメ) どがあります.また,公式ページ[1]から日本 語の有用なドキュメントが辿れます.

本稿執筆時点 (2003年6月) でのRuby のバージョ ンは,安定版が1.6.8であり,次期安定版である1.8.0 のリリースを準備中という状況です.

②はメソッド先頭からのオフセットとバイトコード、およびそのバイトコードに関する付加情報の並びとなっています・バイトコードとは仮想マシンであるJVMの命令のことです・たとえば、addメソッドが最初に実行する命令はiconst\_0で、その次にistore\_2を実行する、という意味になります・

#### クラスファイルの構造

では, クラスファイルの構造はどのようになっているのでしょうか. JVM 仕様 [3] によると, C ライクな文法で表現した場合, リスト2 のようになります.

リスト中のu1,u2,u4はそれぞれビッグエンディアンの符号無しの1,2,4バイト整数を表します.細かいところ,面白くないところは省略しますので,詳細はJVM仕様を確認してください.ちなみに,インタフェースもクラスファイルによって表現されます.インタフェースの場合,access\_flagsにその旨が記録されます.

まず, magic とはこのファイルがクラスファイルであることを示す数値 0xCAFEBABE という値です.おしゃれですね.

コンパイル時に確定する定数や文字列などは, コン

図3 javapの実行結果

```
$ javap -c Counter
Compiled from RavaTest.java

class Counter extends java.lang.Object {
    public static void main(java.lang.StringEl);
    Counter();
    public int add(int);
    public int get();
}

Method int add(int)
    0 iconst_0
    1 istore_2
    2 goto 18
    5 aload_0
    6 dup
    7 getfield #9 <Field int counter>
10 iconst_1
11 iadd
12 putfield #9 <Field int counter>
15 iinc 2 1
18 iload_2
19 iload_1
20 if_icmplt 5
23 aload_0
24 getfield #9 <Field int counter>
27 ireturn
```

スタントプール (constatn\_pool) に詰め込まれます. クラスファイルではいろいろなところからこの情報を利用します. たとえば, そのクラスと基底クラスの情報を格納する this\_class, super\_classは, 実際はコンスタントプールのインデックスとして表現されます.

methodsにはメソッドの数だけエントリがあり、それぞれのエントリが1つのメソッドの情報を格納しています、メソッドの情報にはそのメソッドがどのようなパイトコード列を持つか、例外が発生した場合どうするか、などの情報が格納されています。

fieldsやinterfacesも見ればわかりますね. どんなフィールドがあるか,どんなインタフェースが 定義されているか,という情報です.

## クラスファイルのロード

やっとRava の話です. Rava はもちろん実行時にクラスファイルをロードします.

JVM 仕様としては, クラスファイルが正しいものであるかどうか, 検査 (Verify) しなければならず, その項目も決められています. これは悪意のあるものがクラスファイルを改竄してJVM が不正使用されるようなことを防ぐための機構です.

しかし, Ravaではそのようなチェックを行っておりません.扱うクラスファイルはすべて信頼できるものとして実行します.これはネットワーク越しでクラスファイルをロードするようなことがないという前提でRavaを開発していたからです注3.

リスト2 クラスファイル構造体(C言語ライクな表現)

```
ClassFile {
    u4 magic;
    u2 minor_version;
    u2 constant_pool_count;
    cp_info constant_pool_count=1];
    u2 access_flags;
    u2 this_class;
    u2 super_class;
    u2 super_class;
    u2 interfaces_count;
    u2 interfaces_count;
    u2 fields_count;
    field_info fields[fields_count];
    u2 methods_count;
    method_info methods[methods_count];
    u2 attributes_count;
    attribute_info attributes[attributes_count];
```

注3:一番の理由は面倒くさかったから,ということなんですが

そのため、クラスロード時の一大イベントである検証のプロセスはすっ飛ばして、クラスをロードします、クラスファイルはパイナリファイルということは先ほど述べました、Rubyなどのスクリプト言語はパイナリファイルを扱うのが苦手である、というイメージがあるかもしれませんが、それがそうでもありません、Rubyの場合、String::unpackやArray::packメソッドを利用してパイナリデータを簡単に扱えます。

また,クラスファイルをロードしたときに基底クラスがロードされていなかった場合,基底クラスもロードします.これによって,クラスの派生関係をRava内部で正しく把握することができます.

## 解釈実行

クラスファイルを読み込んだら,それを解釈して実行します.基本的には例外を無視すれば,次のようなループを実行します<sup>注4</sup>.

do{

PC が指す場所の命令を取得;

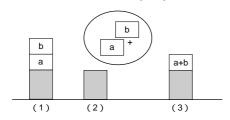
取得した命令の操作を実行;

}while(まだすることが残されているか?);

PC は実行すべき命令位置を示すもの(Program Counter)です.JVM の主な仕事はこれだけです. JVM って簡単でしょ?<sup>注5</sup>

本節では以降,JVMの実行に不可欠な要素を概観

図4 スタックマシンで足し算(a+b)を行う様子



- (1) 値a, bをスタックに積む
- (2) 値a , bをスタックから取り出し , a+bを計算する
- (3) 値a+bをスタックに積む

していきます.

## スタックマシン

JVM はスタックマシンとして実現されています.スタックマシンとは,演算する対象や演算した結果をスタックに保存するアーキテクチャです.たとえば足し算をする場合,スタックから2つの値を取り出し,それらを足し合わせた結果をスタックに積みます(図4).

JVM ではこのような計算をオペランドスタックとい うスタックを用いて実行します.

Rava ではオペランドスタックをRuby の配列で実現 しました. Ruby の配列にはスタック操作をするため の十分な機能がサポートされているためです<sup>注6</sup>.

## 型とデータ表現

Java の型はおおまかに言って , プリミティブ型とリファレンス型に分かれます . 要するに数値などのデータか , オブジェクトか , ということです .

JVM上では型に応じた命令セットを持ちます.ただし,byte,short,char型については,だいたいint型と同じ命令によって計算されます.

Ravaではプリミティブ型はそのままRubyの数値型で表します.リファレンス型,たとえばJavaのオブジェクトは,フィールド領域とそのクラスへのリファレンスを持つRubyオブジェクトとして表現しました.フィールド領域はフィールド名をキーとするハッシュによって表現しました.

JVM の命令セット

JVM では,200もの命令が規定されています.これらを詳細に述べていくと,誌面がすぐに尽きてしまうので,どのような命令セットであるか,JVM 仕様にある命令セットのカテゴリを紹介します.

- ロード命令とストア命令
- 算術命令

注4:JVM仕様の記述と若干変更してあります.

注5: これ以外の主でない部分が面倒なんですけどね

注6: たとえばJavaの配列は固定長ですが, Rubyの配列は可変長です.要するにRubyの配列がjava.util.ArrayListのようなものということです(もっと高機能ですが). 効率は犠牲になりますが, プログラミングは楽になります.

- 型変換命令
- オブジェクト生成と操作
- オペランドスタック管理命令
- 制御の移行命令
- メソッド起動とリターン命令
- 例外のスロー
- finally の実装
- 同期化

各命令は1バイトのオペコードに続いて,引数,あるいは命令が用いるデータとなるオペランドを0個以上続けたものから構成されます.中には可変長のオペランドを持つ命令もあり,なかなか凶悪です.

余談ですが、仮想じゃない機械の命令セットは、命令のビットパターンで、この命令はどういう処理をする、ということがわかるそうですね、命令デコードが簡単になるからだと思います、が、JVMの命令セットはただ機能ごとに番号を振っていっただけなのでそのようになっていません、以前JVMをチップ上に載せようとしていた人が嘆いていました。

## メソッド起動

メソッドを起動すると,各スレッドが1つずつ持つ JVM スタック上に"フレーム"を作ります.フレーム を利用してJVM はメソッドを実行します.

フレームにはそのメソッドで利用する "ローカル変数領域" とメソッド実行用の"オペランドスタック", そしてそのメソッドのクラスに対する実行時コンスタントプールへの参照が含まれます.

メソッドが終了(リターン命令による終了や例外発生による途中終了)したら,JVM スタックからそのメソッドが構築したフレームが取り除かれ,呼び出し元メソッドの実行が再開されます.値を返す場合,呼び出し元メソッドのオペランドスタックへ値を積みます.

Ravaでは具体的に図5のようなフレームを作ることにしました.まずローカル変数領域があり、その上に呼び出し元メソッドのフレームを復帰するためのフレーム情報を積み、その上をオペランドスタックが1つの配列で表現されており、フレーム、そしてオペランドスタ

ックをその配列の一部として実現しました.

メソッドが終了すると,フレーム情報を利用して呼び出し元メソッドのフレームを復元します.

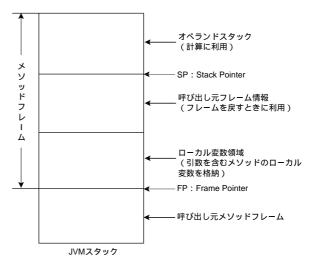
#### ネイティブメソッド

JVM 命令セットには,入出力(環境依存の操作) やスレッドの制御(VMの操作)などを行うための命 令はありません.これらを実現するためにJava仕様, およびJVMにはネイティブメソッドがあります.

ネイティブメソッドとはJava以外の言語によって記述されたメソッドを実行するためのもので、たとえば C言語で記述したメソッドによって入出力を実現した リ、JVMのスレッド生成を行います。Javaプログラム中、メソッド名の前にnativeとして記述すれば、コンパイラはそれをネイティブメソッドとして認識し、クラスファイルにもそのように記述されます。ネイティブメソッドをどのようにJVMとパインドするか、などはOSのローダなどとの兼ね合いから実装依存となっています。また、ネイティブメソッドをC言語などでどのように記述すればよいかを定めたAPIであるJNI (Java Native Interface)[9]という仕様もあります。

JVM 仕様では, どのようにネイティブメソッドを記

図5 Ravaのメソッドフレーム



## ☞ 参考文献

[ 9 ] Java Native Interface: http://java.sun.com/j2se/1.4/ja/ docs/ja/guide/jni/ 述し実行させるかということを規定していないため、 RavaではネイティブメソッドをRubyで記述できるようにしました. つまり, Java プログラム中でネイティブメソッドが起動されたとき, 対応するRuby メソッドを起動するのです.

このおかげで従来はC言語などで記述しなければいけなかったネイティブメソッドがRubyで簡単に記述できるようになりました.しかし逆に,ネイティブメソッドに関する過去の資産が使えません.そのため,たとえばウィンドウシステムを利用するようなJava APIはRavaではサポートしてません.

ちなみにリスト1のrj.outメソッドはネイティブ メソッドを利用する出力メソッドです。

#### 例外処理

Javaには例外処理機構があります。try/catchで挟むあれです。try節で発生した例外は,その例外がcatch節でフックされていれば,そのcatch節が実行されます。また,finally節がある場合,例外が起こったか起こらなかったかに関わらずfinally節が実行されます。catchされない例外は,呼び出し元メソッドへ例外を伝播します。

さて、JVMではこれを"例外テーブル"(表1)というものを利用して実現します。例外テーブルは各メソッドそれぞれが保持しています。例外テーブルによって、"どこで例外が起こったか"、"どの例外が発生したか"という情報をキーとして、"どのような処理を行う"という情報を引き出し、例外処理機構を実現します。

表1 例外テーブルの例

start_pc	end_pc	catch_type	handler_pc
10	20	Α	100
10	20	В	200

オフセット10~20を実行中に例外Aが発生したら処理を100 に,Bなら200に移す

## ☞ 参考文献

[10]『Ruby ソースコード完全解説』 / 青木峰郎著 / まつもとゆきひろ監修 / インプレス 2002

Ravaではこの動作をRubyの例外を利用して実現しました.つまり,解釈実行のループ中でJavaの例外が発生した場合,Rubyの例外を発生させてしまいます.ループから出た時点ですべてのRubyの例外を補足し,現在実行中のJavaメソッドの例外テープルを検索します.もし該当エントリがあればその情報を元にJVMの実行を再開します.なければメソッドフレームをさかのぼって検索を続けます.

## GCやスレッド

ガーベッジコレクション (GC) は不要なメモリをシステムが勝手に回収し,再利用可能にする機能です.スレッドは実行する命令流を仮想化し,ソフトウェアで複数制御できるようにしたものです.

なぜこの2つを並べて書いているかというと, Rava ではこれらの実装をRubyのGC およびスレッド機能でそのまま実現しているためです.たとえば, Javaのオブジェクトをnewすると, Rava はRubyのオブジェクトを生成します.そのJavaオブジェクトが必要なくなったとき, RubyのGC 機能がいいようにしてくれます.楽です.

スレッドは, Java のスレッドが生成されたとき, Ruby のスレッドを生成して, その上でバイトコードの解釈実行を行います. 楽です.

ただし、これらの機能はまだ厳密にJVM 仕様を満足しておらず、たとえばJava のファイナライズやスレッドの同期などはRava では未着手です<sup>注7</sup>・特にスレッドの同期に対応していないのはスレッドプログラミングにとっては致命的なので、Rava ではスレッドはなんちゃってサポートです。

では、RubyでGCやスレッドをどのように実現しているか、そもそもRubyインタープリタはどのように動いているのか、という詳細は文献[10]を参照してください。

## もっと速く!

前節までで説明した範囲を実装することで, Rava をとりあえず動かすことができました.しかし,実行

注7: ごめんなさい. 面倒くさかったんです.

速度が涙が出るほど遅いものになってしまいました.

これは、Ravaがスタックマシンの挙動を非常に素直に、悪く言えば馬鹿正直に実行しているためです。なんとかならないものでしょうか。

世の中ではさまざまなJVMの高速化手法が考案されています.そこで,それらを少し紹介してみます.

JIT (Just-In-Time) コンパイラによる高速化は,実行時にパイトコードをコンパイルする手法です.一般的に,パイトコードを実行環境の機械語列へ変換することでより高速な実行を可能にします.ただし,メモリ使用量が非常に増えます.ほとんどのJVMはこの機構により高速化を行っているようですが,組み込みJVMなどでは利用可能リソースの関係からこれの適用が難しい場合があるようです.

AOT (Ahead-On-Time) コンパイラは,実行時にコンパイルを行うJIT コンパイラと違い,実行前にさっさとコンパイルをやっておこうというものです.実行前にコンパイルを終わらせてしまうため,コンパイル時間を縮小できるという利点があります.

## 静的な言語,動的な言語

Rubyでは、eval メソッドがあったり、クラスを実行時にいじったり、メソッドを上書きできたりと、動的で柔軟な言語です、Java はそのような点で静的な言語と言えると思います。

動的な言語では、プログラミングの幅が広がります. たとえば、本稿で述べた高速化の試みでは、動的にメ ソッドをクラスへ追加するという実装を行いました.

しかし、すべてが実行時まで決まらないため、Ruby は非常に最適化がしにくい言語になっています。たとえば、1+1というRuby プログラムを、パースした時点で2とすることが出来ません。これは、足し算を行うときに呼ばれる Fixnum::+というメソッドの挙動が実行時に変更される可能性があるためです。

Javaは, Rubyに比べるとコンパイル時にほとんどのことが確定するため,最適化に適したつくりと言えます.プログラミングのしやすさを取るか,実行効率をとるか,そのパランスの違いがRubyとJavaとの違いの1つなのかな,と思います.

SunのHotSpotVMなどでは、実行時プロファイリングによって何度も実行する部分(HotSpot)のみをコンパイルしようというアプローチをとっています.たとえば初期化のために一度しか呼ばれないメソッドをコンパイルするよりは、何度も実行されるメソッドをコンパイルしたほうが処理速度は向上します.

#### Rava の場合

どうやら、バイトコードをなにかに変換すると実行速度を速くすることができそうです。そこで、実行時にバイトコードをそれと同等の働きをするRubyスクリプトに変換し、それを実行するような機構(JITコンパイラ)を試作してみました。どのバイトコードをコンパイルするかはメソッド起動回数を集計するような簡単なプロファイラを用意し、起動回数によってコンパイルするメソッドを決定するようにしました。

Ruby ではeval メソッドによって実行時にRuby スクリプトを評価 (Evaluate) することができるため、そのような処理がやりやすかったのです.

この試作では,実行速度がインタプリタのままの性能よりも25倍も向上することができました<sup>注8</sup>(詳細はサポートページ<sup>[1]</sup>をご覧ください).

## おわりに

いかがでしたでしょうか. JVM はこんなことをやっているのか, と雰囲気だけでも掴んでいただければ幸いです. ついでに, Ruby に興味を持った,使ってみたい,と思ってもらえるとこちらの思うツボです.

言語処理系をいじるのは楽しいです。ちゃんと動い てくれると本当に嬉しいです。処理系上で動作するプログラミング言語への理解も深まります。Ruby だといろいろ簡単に作れますし、まだまだやめられそうにありません。

最後に,本稿執筆にあたり,お世話になった方々, 相談にのっていただいた方々,そして最後まで読んく れたあなたに,感謝いたします. □