

情報処理の概念

#8 情報・データ・プログラム・バグ

Yutaka Yasuda

ソフトウェアの安全性

- 1999 米国での報告

「国民が脆弱なソフトウェアに頼らざるを得ない状況にある」 PITAC（大統領情報技術諮問委員会）

- 事例

みずほ銀行システムトラブル（2002/4, バグ）

JR 東日本 Suica 改札トラブル（2006/12, バグ）

NTT東西 ひかり電話接続不良（2006/9,10, 複合）

Yahoo! BB 470万人顧客名簿流出（2004/1, 情報管理）

KDDI 400万人ユーザ情報流出（2006/6, 情報管理）

- ウィルス問題（各社, セキュリティホール等）

Software Update

- 事例：Microsoft のソフトウェア・アップデート

ほとんどがトラブルまたは安全上の対策（バグ）

減る気配なし

Software Update (ex. Microsoft)

- MS07-001 : Office の重要な更新 (921585) (2007/01/10)Office Update
- MS07-002 : Excel の重要な更新 (927198) (2007/01/10)Office Update
- MS07-003 : Outlook の重要な更新 (925938) (2007/01/10)Office Update
- MS07-004 : Windows の重要な更新 (929969) (2007/01/10)Microsoft Update
- MS07-005 : Windows の重要な更新 (923723) (2007/02/14)Microsoft Update
- MS07-006 : Windows の重要な更新 (928255) (2007/02/14)Microsoft Update
- MS07-007 : Windows の重要な更新 (927802) (2007/02/14)Microsoft Update
- MS07-008 : Windows の重要な更新 (928843) (2007/02/14)Microsoft Update
- MS07-009 : Windows の重要な更新 (927779) (2007/02/14)Microsoft Update
- MS07-010 : マルウェア対策エンジンの重要な更新 (932135) (2007/02/14)
- MS07-011 : Windows の重要な更新 (926436) (2007/02/14)Microsoft Update
- MS07-012 : Windows の重要な更新 (924667) (2007/02/14)Microsoft Update
- MS07-013 : Windows の重要な更新 (918118) (2007/02/14)Microsoft Update
- MS07-014 : Word の重要な更新 (929434) (2007/02/14)Office Update
- MS07-015 : Office の重要な更新 (932554) (2007/02/14)Office Update
- MS07-016 : Internet Explorer の重要な更新 (928090) (2007/02/14)Microsoft Update

つづく... (資料は2007/3 末現在)

http://www.microsoft.com/japan/security/bulletins/visual_list.msp
Microsoft, Inc. 絵で見るセキュリティ情報

ソフトウェアの安全性

- 2002 米国での試算

「ソフトウェアの低品質によって米国経済は 600 億ドル (7.2兆円)、GDP の 0.6% の損失をこうむっている」
.. NIST (National Institute of Standards Technology)

日本の GDP から推定すれば 3 兆円に相当

- バグ

ソフトウェア障害の 3/4 が既知の初歩的なプログラミングの誤りによる - Eugene H. Spafford

バグ

- なぜバグは発生するのか？無くならないのか？

「注意深く設計すれば良かった」だけか？

- 情報処理システムが本来持つ構造問題として理解していきましょう

「情報処理」について

情報処理

- 情報処理とはなにか？

実体はデータ処理

= コンピュータはデータ自動処理機械である

- なぜデータ処理機械で情報処理が可能なのか？
- 情報とデータの関係に注目

情報

A B C

- 「情報とは？」

これはインクのシミか？それとも A という文字か？

物理で記述できない違いこそ「情報」

- 情報

ものごとの説明（特徴を抽出したもの）

- データ

情報を一定ルールで「記述」したもの

（あるいは「表現」したもの。符号の列。）

データ

- コンピュータは情報とデータを区別する

コンピュータは：情報が持つ意味を解釈しない

人間が：データ処理を情報処理として解釈する

- それを前提に：
情報処理機械とデータ処理機械は等価とする
- コンピュータが情報とデータの役割分担を明確にしたとも言える

具体的なコンピュータの中での
データ処理のすがた

入出力装置

データを取り込んだり、取り出したりする。
データは一般に内部回路で直接扱えない形式をしている。例えばスイッチのon/offや印字結果など。つまり man machine interface が必要。

コンピュータの構成

演算装置

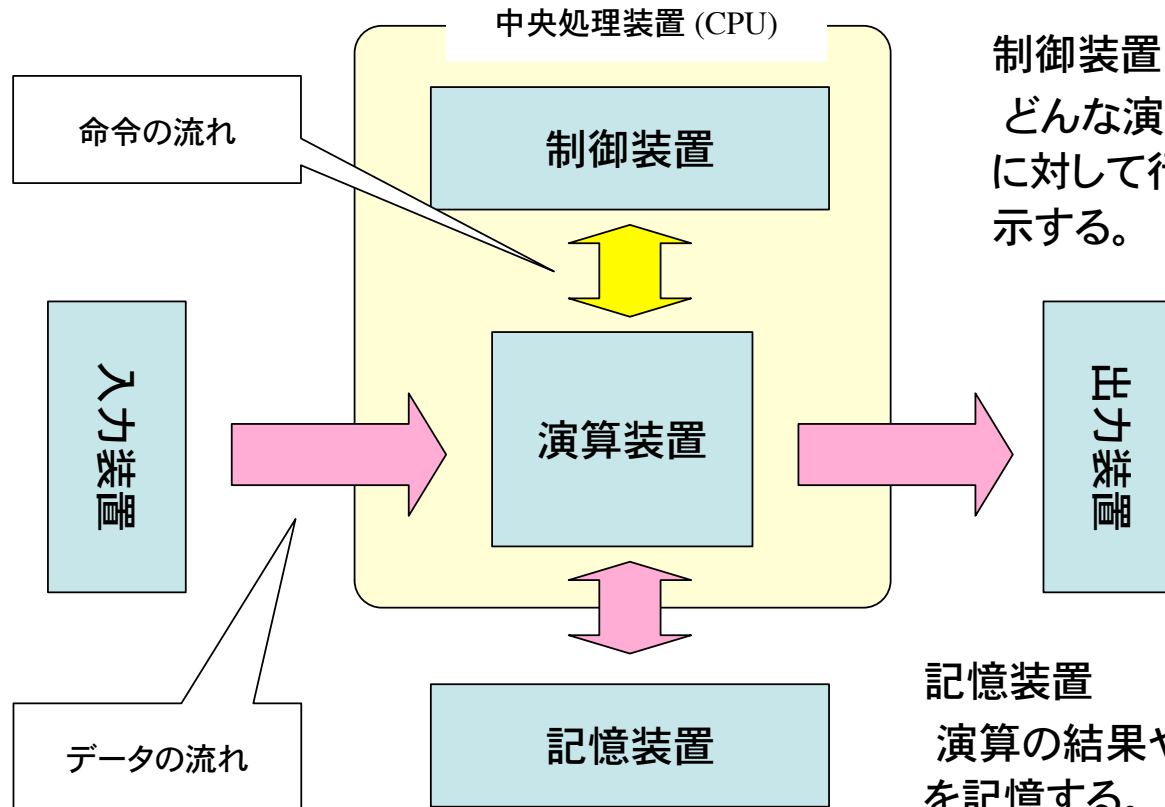
1+1などの単純な演算を行う。
結果はほんの僅かしか記憶できない。

制御装置

どんな演算を、入力・記憶していたどの値に対して行い、どこに出力・記憶するかを指示する。

記憶装置

演算の結果や、次に行うべき演算の指示内容を記憶する。
順序だてた動作のために状態の記憶が必要。
(今どうなったかを基に、次どうするかを決める)

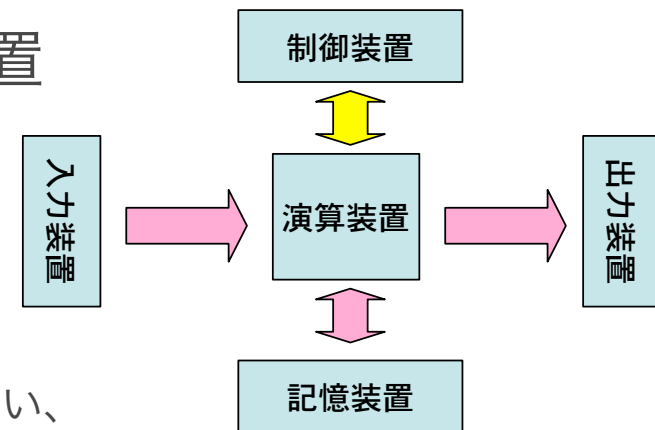


コンピュータの構成要素

- 現代的コンピュータの構成要素

入出力装置 + 演算装置 + 制御装置
+ 記憶装置 (+ 自動化)

データを入力し、
(単純な) 演算をほどこし、
結果や状態を保存しながら、
繰り返しや条件分岐など順序だてた処理を行い、
結果としてデータを出力する



- ノイマン型

1946, ノイマンらが提案した構成モデル

図はノイマン型コンピュータの論理的モデル

自動情報処理機械

- 自動計算機械（電卓的計算機）はできた
情報のデータ化（符号化）
汎用性＝計算機をワープロとして使う
- 自動データ処理機械もできた
手順＝ソフトウェアの導入
- 自動情報処理機械はどのようにして？
- 現状での結論
「データ処理を情報処理と解釈できる」

ソフトウェアとバグの関係

バグ

- プログラムに含まれる「間違い」
- データは意味をもたない
- コンピュータは意味を扱わない
- バグ

無意味な（矛盾した）処理でも指示通り動作する

例えば「金利と残高を加算する」ところを間違えて
「年齢と金額を加算」させてしまうかもしれない

- なぜ間違えた指示が最後まで？

プログラミング

- 「1から10までの数を足した結果を出せ」
これはコンピュータにとって「難しすぎる」指示
- 手順の明記
「Xを1から10まで変化させ、毎回Yに繰り返し込め」
- プログラムとは何か
目的に対して「何をどう処理するか」を詳述したもの
- 意味の消失
「コンピュータはデータの意味を理解しないのと同様に、プログラムの意味も知らない」
(ただ手順だけを知っている)

手順をどのように書くか

- コンピュータは日本語を理解できない

「Xを1から10まで変化させ、それを毎回Yに繰り返す」のもコンピュータには複雑すぎる

コンピュータが理解できる言語で書き直す必要がある

C言語での例

```
Y=0;  
for (X=1; X<=10; X++) {  
    Y=Y+X;  
}
```

Yははじめ0だと設定している

1から10まで変化させるということを、「1からはじめて10以下の場合は終わりまでの処理を行い、1加算してもう一度繰り返し」という表現で明記している

Yに増え続けるXを足したものを再びYに代入

手順をどのように書くか

- しかしC言語でもハードウェアには直接理解できない
さらに単純な処理に分解しなければ
CPU(例は Motorola 68000)向けのアセンブリ言語で記述

```
    st 0,$1004
    mov 1,%o0
    st %o0,$1000
.LL2:
    ld $1000,%o0
    cmp %o0,11
    bgt .LL3
    ld $1004,%o0
    ld $1000,%o1
    add %o0,%o1,%o0
    st %o0,$1004
    ld $1000,%o1
    add %o1,1,%o0
    mov %o0,%o1
```

C言語での例

```
Y=0;
for (X=1;X<=10;X++) {
    Y=Y+X;
}
```

手順をどのように書くか

- アセンブリ言語でもまだ直接は理解できない。
更に機械向けの言語＝機械語に直さなくては

(以下は機械語の一部)

```
7400 5f5f 4354 4f52 5f4c 4953 545f 5f00
5f65 6e76 6972 6f6e 005f 656e 6400 5f47
4c4f 4241 4c5f 4f46 4653 4554 5f54 4142
4c45 5f00 6174 6578 6974 0065 7869 7400
....
```

情報処理から機械語へ

- 処理からプログラムまでの変換過程
 1. 「1から10までの合計」という処理を
 2. どのようにして計算するかという手順に分解
 3. それをプログラミング言語で記述し
 4. 機械語まで変換
- 処理モデルから手順列への変換
 - この作業をプログラミングと呼ぶ
- プログラミング言語から機械語への変換
 - ハードウェアに解釈可能な動作指示へ

二つの変換過程

人間側

コンピュータ側

1から10までの数を
足した結果を得る

人間が変換
(プログラミング)

この時点で意味が消失
して手順だけが残る

つまりバグが含まれていても検証
できない

```
Y=0;  
for (X=1;X<=10;X++) {  
    Y=Y+X;  
}
```

機械が変換

ここで実行されてはじめてバ
グが見つかる

02af93e8f
37de76e0
4e3a2...

失われる意味

- プログラミングの過程で、プログラムから意味は失われる
そのプログラムの目的を知るのはプログラマだけ
- コンピュータは情報ではなくデータだけを扱う
データが表現する情報は入出力の前後にいる人間が扱う
- ソフトウェアは作業から意味を抜いた手順だけを扱う
その意味（内容）は結果を受け取る人間だけが知る
- バグ（意図しないプログラムの振る舞いをひきおこす不具合）が発生する根本的要因のひとつ

大規模システム開発の問題

Brooks に学ぶ

まとめと対策

- なぜトラブル・バグはなくなるらないか？
- 情報とデータの相違
 - コンピュータは情報でなくデータを扱う
 - 情報処理機械≠自動データ処理機械
- ソフトウェア工学の価値
 - 大規模システム開発における問題点
- 脆弱なソフトウェア基盤からの脱却を

Y2K 問題を思い出せ (or 調べてみよ)

- 坂東俊矢 (京都産業大学法務研究科)
「IT2001 なにが問題か」から

「どこにあるかも知らないコンピュータの誤動作によるライフラインの断絶に備えて、半信半疑で風呂に水をため、当面の食料品を購入したりもしたが最後まで具体的な情報は消費者には伝わってこなかった」

「消費者は情報社会が滑稽なほど不完全であって、それが自らの生活に直接の影響を与えるものであることだけは理解できたに違いない」

- 参照：首相官邸「コンピュータ西暦2000年問題」
<http://www.kantei.go.jp/jp/pc2000/>