

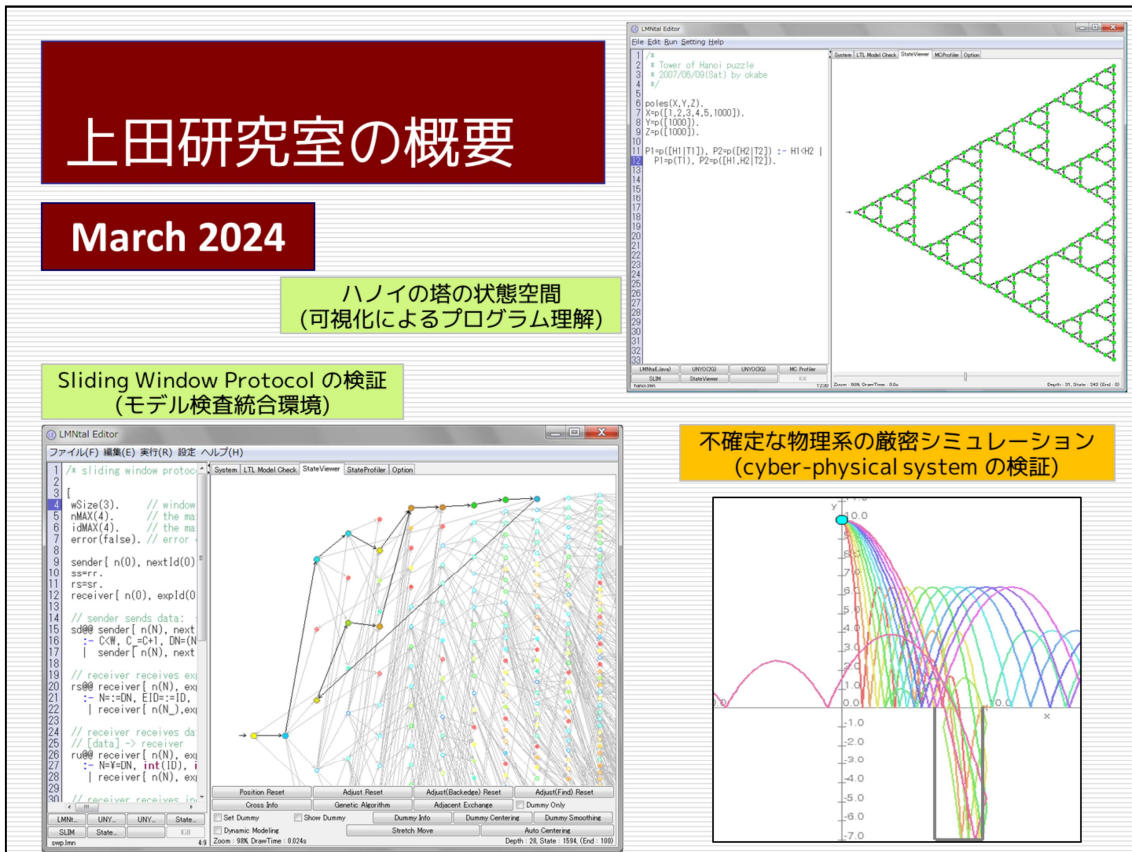
上田研究室の概要

March 2024

ハノイの塔の状態空間
(可視化によるプログラム理解)

Sliding Window Protocol の検証
(モデル検査統合環境)

不確かな物理系の厳密シミュレーション
(cyber-physical system の検証)



こんにちは。

さっそくですが、この表紙のスライドは、上田研でゼロから開発したプログラミング言語とビジュアル開発環境を用いたプログラム実行例を示しています。

右上は皆さんご存知の「ハノイの塔」のプログラム（字が小さいですが）とそれが生成する状態空間です。操作ルールは再帰を使わずに1行（横幅の関係で折り返されてます）で書け、ボタン一発でフラクタル的な状態空間が図示できます。
(ハノイの塔の状態空間図を見たことがありますか?)

左下はコンピュータネットワークの授業で学ぶTCPで使われている sliding window プロトコルの検証を可視化したもので、通信路が信頼できない場合にパケット再送が無限回繰り返される例を状態遷移図からハイライトしています。

この二つのビジュアル環境は実は同じものですが、皆さんの先輩が4年生の時に着想して構築したツールが発展してきたものです。

右下は物理系のシミュレーションで、穴の開いた地面を弾むボールの軌道を可視化しています。普通のシミュレーションとまったく違って、浮動小数点計算を使わずに厳密な計算を行い、しかもボールの初速度によって変わるさまざまな挙動を自動的にかつ一挙に示しています。

どちらも、世界的にオンリーワンのツールです。

コンピュータサイエンスの核心に挑む

「難しい機能を実現する」「高性能で」「正しさが保証された」ソフトウェアに向けた**開拓研究**をしています

1. 先端的プログラミング言語のメカニズム設計

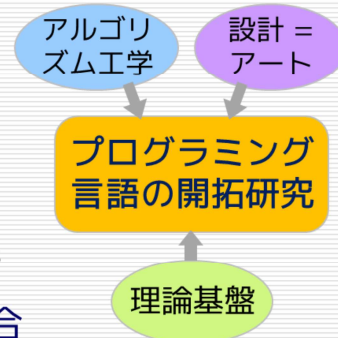
(理論+デザイン+アルゴリズム) から**処理系開発**まで

既存の言語が苦手な分野は多数！

- ★ 複雑なデータ構造を簡明に操作
- ★ プロセスやプロセッサの分業
- ★ 連続量 (実数) を正しく扱う
- ★ 高機能3D印刷

2. 検証つきソフトウェアの構築技術

- ★ 言語処理系とプログラム検証系の統合
- ★ 高性能並列検証技術



2

上田研はこのようなオンリーワンの先端ツールを実際に開発しながら、コンピュータサイエンスの核心を探求することをモットーにしています。

コンピュータの万能性はプログラムのおかげですから、コンピュータサイエンスの核心はプログラムのサイエンスであるといって過言ではありません。プログラムは単なる手段にすぎないと思っていると、情報社会全体が砂上の楼閣になります。誰かが核心を深く追究し、先駆的な試みに挑戦する必要があります。

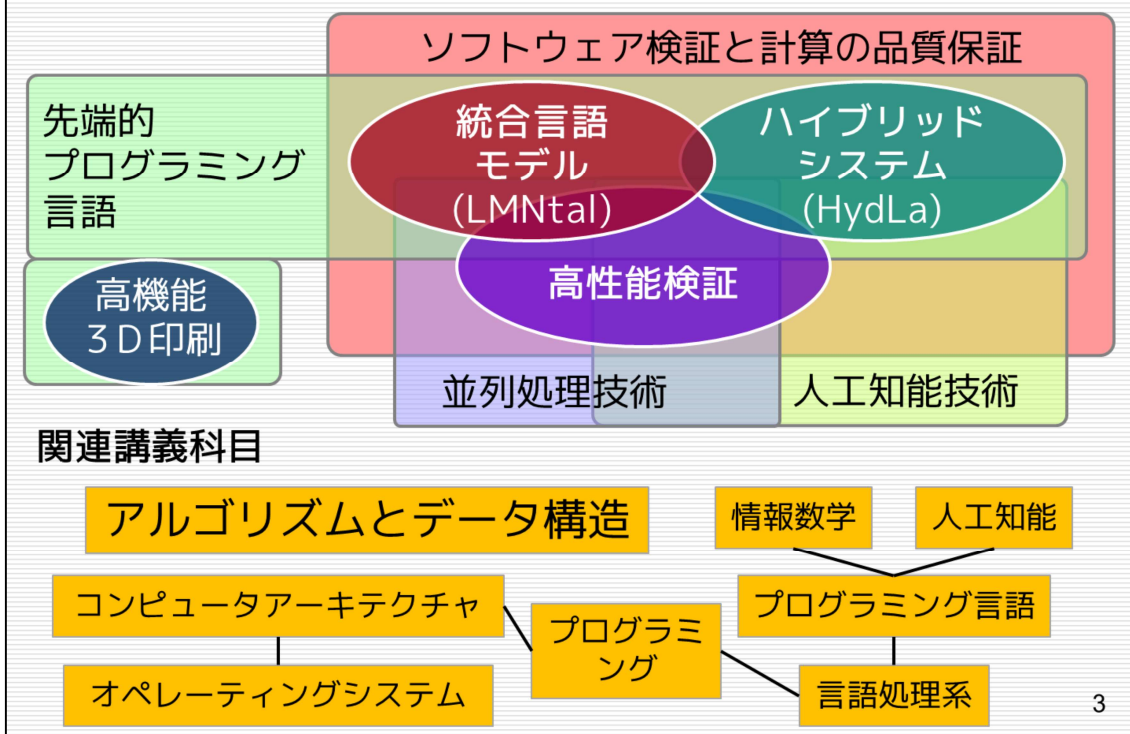
我々は、難しいと考えられている機能を容易に実現するソフトウェア、ハードウェア性能を限界まで引き出すソフトウェア、プログラムが書けたときから正しさが保証されるようなソフトウェア、などをいかにプログラムするのか、という問題意識で技術開発に取り組んでいます。

この目標を実現するために、我々は、先端的なプログラミング言語が備えるべきメカニズムの基本設計から行います。言語のメカニズム設計は、数学的理論、アート、アルゴリズム工学、の三者が交わる非常に楽しい仕事です。

基本設計ができれば処理系構築です。このような作業を、多数のプロセッサが分業するソフトウェア、アルゴリズム未開拓領域のソフトウェア、連続量を正しく扱うソフトウェア、高機能3D印刷のためのソフトウェア、など多様な分野で展開しています。

もうひとつの観点は、正しさが検証されたソフトウェアの構築技術です。社会基盤で用いるソフトウェアでは、正当性の検証は特に重要です。このニーズにこたえるために、我々は言語処理系とプログラム検証系との統合や、並列コンピュータをもちいた高性能検証技術に取り組んでいます。

研究グループと相互関係



このスライドは、我々の研究グループと相互関係を表しています。楕円がプロジェクト、四角が技術を表しています。

いまは二つの言語、LMNtal と HydLa を核としたプロジェクトを中心に展開してきて（赤と緑の楕円）、3Dプリンタを格段に高機能化するための言語の開発にも着手したところです（濃緑の楕円）。並列マシンを使ったソフトウェア検証を目指す高性能検証（紫色の楕円）も、熱意ある人が推進してきました。

ほとんどのプロジェクトはソフトウェア検証と計算の品質保証技術（薄赤の四角）を重要目標としています。さらに並列処理技術（薄青の四角）と人工知能技術（黄緑の四角）も重要な技術となっています。

これらの研究に関連する講義科目は、スライドに書いてあるような情報系のコア科目です。アルゴリズムとデータ構造は特に大事なのでちょっと大きな字にしました。

これらのコア科目に興味のある人は、上田研に特に向いていると思います。実は、トップIT企業が最もほしがるのが、何よりもこれらのコア科目をしっかり勉強した人たちだったりします。

先端的プログラミング言語の設計と実装

- ◆ ポインタを駆使した複雑な計算から λ 計算に至る多様な計算概念を簡明に統合する言語を開拓
 - **統合プログラミング言語 LMNtal (Github)**
 - ★ 形式意味論からコンパイラ (in Java), 並列検証系 (in C++), 可視化統合環境まで, ゼロから共同開発
 - ★ 皆さんの先輩の大活躍で十数万行規模に発展 (毎年温泉で開発合宿!)
 - ◆ 連続と離散をきちんと統合する言語を開拓
 - **ハイブリッド制約言語 HydLa (Github)**
 - ★ 人工知能, プログラミング言語, 数値計算の総力戦
 - ★ 例年学生と国際共同研究出張 (+ 開発合宿)

世界的にも数少ない開発拠点!

4

研究室の中心テーマはプログラミング言語ですが, 上田研の最大の特徴は, 既存のプログラミング言語にとどまらず, プログラミングのための真に新たなコンセプト (考え方) を発信し, さらに実現している点だと思えます。

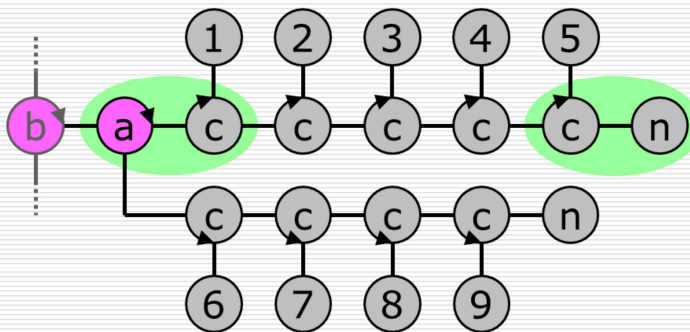
このスライドではそのような二つの言語を紹介します。

最初の言語 LMNtal (elementalと読みます) は単純だけど多才な言語で, C言語で書くとぐじゃぐじゃになってしまいそうなポインタを駆使した計算から, λ 計算のようなフォーマルな計算モデルまで, 多様な計算概念を統一的かつ簡明に統合することを目指しています。皆さんの先輩が力を合わせて, 細部にわたる言語設計から, コンパイラ (Java), 実行・検証系 (C++), 可視化統合環境 (Java/Scala) まで, ゼロから開発してきました。Githubで見ることができます。

もう一つの言語 HydLa は, コンピュータや通常のプログラムが苦手とする真の連続値 (実数値) を, 離散値と統合的に扱うことへの挑戦です。このような言語や処理系を作ろうとすると, プログラミング技術だけでなく, 人工知能や数学もフル活用する必要があります。開発拠点が少ない中, 海外パートナーの大学に学生と出かけて共同研究をしたりします。

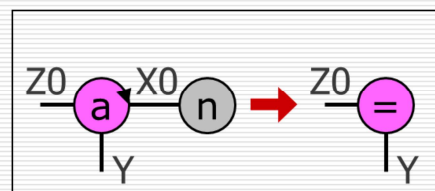
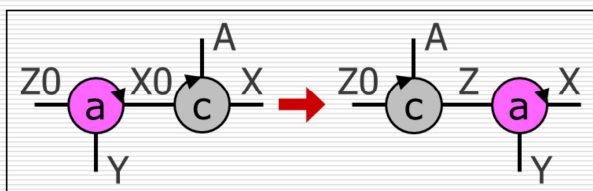
このようなお話をすると, 実世界から浮いた話をしているように聞こえるかもしれませんが, 実はそうではありません。処理系は C++ や Java などのメインストリーム言語で書かれており, 力のある先輩のリーダーシップで途切れることなく進化発展を続けています。

LMNtal: towards a “Turing Machine” of the 21st century



言語と処理系の特徴

- 計算はグラフ構造の変形（単純・強力でポインタより安全）
- スレッドより簡明で安全な並行処理
- システム検証用ツール（モデル検査器）へ発展進化



この2つのルールがリスト連結プログラム

5

これが一つめの言語 LMNtal の説明図です。

LMNtal では、データ構造はグラフ、計算はグラフの書き換えとして表現します。21世紀に入りたての頃は「21世紀のチューリングマシンを目指す」がスローガンでした（スライドタイトル参照）。

プログラミングBの授業で OCaml でリスト連結プログラムを書きましたが、スライドの下の方の二つの四角が LMNtal によるリスト連結プログラムです（テキストでも書けます）。

これをスライドの上の方のグラフに作用させると、整数1～5のリストと6～9のリストが連結されて、1～9が並んだリストができます。

（本当は Powerpoint の図がアニメーションになっています。）

これだけだとトイプログラムにしか見えないかもしれませんが、実はこいつが「大化け」して、ソフトウェア検証のための言語+ツール（4年科目で扱うモデル検査器）として発展進化してきました。

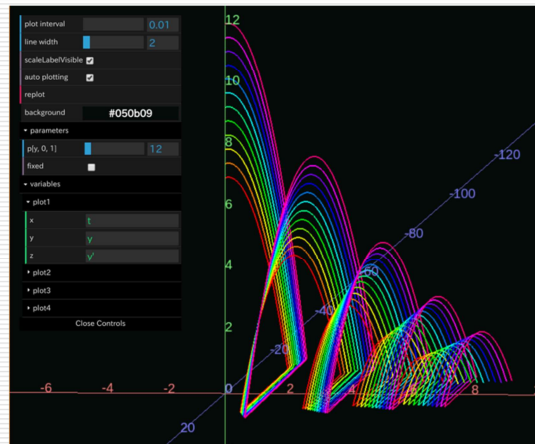
単に実装を改良するだけでなく、新たな言語機能の提案も学生主導で続々と行われ、優れたものはメインストリームに統合されてきています。

HydLa: ハイブリッド制約モデリング言語

- ◆ 連続と離散が共存する**サイバーフィジカルシステム**のための高水準モデリング言語とその高信頼処理系
- ◆ 論理・制約等の**人工知能基盤技術**をフル活用
 - ★ プログラミング言語・人工知能・数学の接点
 - ★ 実数計算と条件判断の正しい共存が挑戦的課題

```
INIT    ⇔ 7 < y < 12 & y' = 0.  
FALL    ⇔ □ (y'' = -10).  
BOUNCE ⇔ □ (y = 0 ⇒ y' = -4/5 * y'-).  
  
INIT, (FALL << BOUNCE).
```

弾むボールのHydLaプログラムと
webHydLa上での実行結果



これは二つめの言語 HydLa の説明です。

サイバーフィジカルシステムという用語を聞いたことがありますか。コンピュータと物理世界からなるシステムで、サーモスタットから自動運転車まで、電子機器やコンピュータと実世界とが相互作用する系はすべて含まれます。コンピュータは基本的に離散で物理世界は基本的に連続なので、離散と連続が共存するのが特徴です。

このような系をモデリングしてシミュレーション・検証するために、HydLa という言語とその処理系を作ってきました。左下が弾むボールの HydLa プログラムの例で、微分方程式を含む論理式を書くだけでシミュレーションしてくれます。さらに、初期条件が具体的に分かっていなくてもシミュレーションができるのが特徴です。右下が、統合環境 webHydLa で可視化した実行結果です。

上のようなふつうのシミュレータにない機能を実現するために、論理や制約などの人工知能基盤技術をフル活用しています。

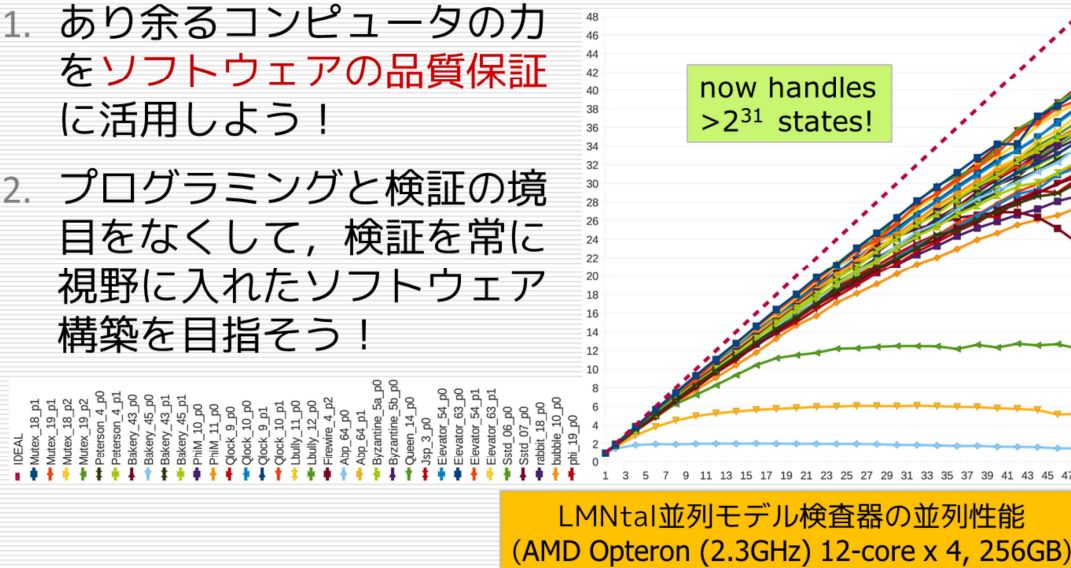
実数計算と条件判断は相性が悪くて、きちんと考えないと、シューティングゲームを作っても条件判断がちゃんとできなくて、弾丸が標的をすり抜けてしまったりします。この相性の悪い二つを数学的にきちんと共存させるために多くの工夫をこらしています。

システムは Github で公開を始めています。

ソフトウェアの高性能並列検証

以下の2つの動機の具現に取り組んでいます

1. あり余るコンピュータの力をソフトウェアの品質保証に活用しよう！
2. プログラミングと検証の境目をなくして、検証を常に視野に入れたソフトウェア構築を目指そう！



M. Gocho, T. Hori and K. Ueda, *Computer Software*, **28**(4), 2011 7

我々の身近にあるコンピュータのほとんどはマルチコアかメニーコアですが、その計算パワーを自分自身で活かしている人、活かせる人はごくわずかというのが現状です。あり余る計算パワーの使い方として、最近ではディープラーニングが主流になってきましたが、我々はソフトウェアの品質保証に活用するという、社会基盤的に重要な方向で研究を進めています。

このスライドは、ソフトウェア検証の代表技術であるモデル検査ツールの並列効果を示すものです。ある学年の先輩がモデル検査器を作り、次の学年の先輩がプログラムを並列化して数々の工夫を施したところ、48コアのマシンで見事な並列効果が出ています。

このツールは、いまでは2の31乗、つまり数十億状態の状態空間を扱うことができるようになっています。

プログラミング言語研究の魅力

1. 言語はコンピュータサイエンス（技術と文化）の『かなめ』
 - ★ 情報技術全体が見渡せる
 - ★ 良いアイデアやソフトウェアは普遍的な価値が驚くほど長持ち
2. マルチリンガルは強くかつ柔軟な技術者への道
3. 処理系開発経験（in C/C++/Java/...）を通じてアルゴリズムとデータ構造，アーキテクチャ，OSに強くなれる
 - ★ トップ企業（IBM, Google, ...）が一番欲しがるのは上記コア科目の技術を使いこなせる人材！

We're working here!

OCaml / Haskell /
Lisp / Prolog /...

C++ / Rust /
Julia /...

C / Java /
Python /...

8

プログラミング言語に興味を持つことの利点や強みは何でしょうか？

一つは、言語はコンピュータサイエンスの技術や文化の「かなめ」となっていることです。「かなめ」を押さえておくと、情報技術全体が見渡せます。また「かなめ」の中にあるアイデアや、「かなめ」を実現するソフトウェアつまり言語処理系は、他のアプリケーションソフトとは大きく異なるタイムスケールで、普遍的な価値を保つことができます。

次に勧めたいのが、いろいろな言語に興味をもつことです。ただし、似た言語をたくさん知っていてもダメで、考え方の異なる言語にふれることが必要です。右上の図にいくつかの言語が並んでいますが、おそらく、階層の上の方の言語を知っている人は下の方の言語も知っているし使えるけど、逆は必ずしも成り立たないと思います。これが、マルチリンガルな技術者の強さです。

我々はプログラミング言語の基礎理論から実践までカバーしますが、実践で一番重要なのは処理系作りです。これは普通のプログラミングとは規模も難度も違いますが、卒論生でもその一部にかかわることが十分可能です。

いずれにせよ、言語処理系はアルゴリズムとデータ構造、コンピュータアーキテクチャ、OSなどの知識と技術の総動員なので、プロジェクトにかかわることでこれらの分野の実力が自然に上がります。前にも述べたように、トップ企業が一番欲しがっているのが、最新技術や動向よりも、これらのコア科目の技術を使いこなせる人材です。

研究室のポリシー

詳しくはオープンハウスや
研究室Q&A集で！

- ◆ **地頭を鍛えよう** — 本物の問題解決力
 - ★ 研究室によく来て環境とチャンスを最大限利用しよう
- ◆ **実機・実システム主義** — 開拓研究の体験を積む
 - ★ 有数の開発拠点で only one のソフトウェアを目指して
並列マシン, コードベース, それを支える理論と格闘
- ◆ **根本原理を重視** — 10年先の情報技術を先導できるように
 - ★ ソフトウェアの科学技術の現状に強い問題意識をもち,
新たなパラダイムと一緒に開拓したい人を歓迎します

プログラミングは、好きならば理想、嫌いでなければOK。
共同開発が多いので**協調性**は必要です。あとは**日本語能力+**
機転と頓知+スタミナ。システムソフトウェアや並列ソフト
ウェアは難度が高いですが、卒論で数千行、修士で1~2万行
のシステムを構築公開する人が稀ではありません。

9

どの研究室にも研究の基本ポリシーというものがあると思いますが、我々のポリシーは以下の通りです。

1. 見たことのない新たな問題に対して、多面的に考えて解を見出す力、すなわち地頭（じあたま）を鍛えましょう。AIに一番苦手そうな部分ですね。これには、研究室に良く来て、各研究室特有の環境（専門書、優秀な先輩、おもちゃ、調理設備、その他なんでも）を最大限活用してほしいと思います。
2. 我々は論文を読んだり書いたりするにとどまらず、実マシン上で動く本格的なシステムを相手にします。学生主導でシステムソフトウェアを開発する経験とノウハウは世界有数で、作業のための並列マシンも、研究の土台となるコードベースもそろっています。
3. 研究室で身につけた知識やスキルがすぐ陳腐化しないように、ものごとの根本原理を重視します。根本原理はいったん身につけると一生ものになり、10年先の情報技術を先導する力になります。私見では、いまのソフトウェア技術はまだハイテクと呼べるものではありません。ソフトウェアの科学技術の現状に問題意識をもち、新たなパラダイムと一緒に開拓したい人を歓迎します。

最後に、ここまで読んでくれた人は、上田研は相当なプログラミングスキルを求めていると思うかもしれませんが、そうでもありません。プログラミングが好きならば理想ですが、とりあえず嫌いでなければOKです。ただし共同開発が多いので協調性は必要です。プログラミングのスキルよりも、日本語で物事を表現する力とか、機転と頓智とか、スタミナとかが役立つかもしれません。

ふつうのソフトウェアよりも挑戦的なことに取り組みますが、そばにいる人がざりざり作業する様子を見ることで、門前の小僧効果が期待できます。

個別面談の予定, Q&A集

- ◆ 研究室配属もプロジェクト研究も, ほぼ下記の時間帯に相談可能です. 63号館5階05-02, 05-22です.
最新情報は下記ウェブページで確認してください:
<https://www.ueda.info.waseda.ac.jp/~ueda/kenkyuusitu-j.html>
 - ★ 3月18, 19日 11AM~1PM, 2PM~5PM
 - ★ 3月21日 11AM~2PM
- ◆ ほとんどの時間は先輩との相談も可能です.
- ◆ 事前連絡 (ueda@ueda.info.waseda.ac.jp) は歓迎ですが**必須ではありません**. アポなしで直前連絡で来て, 面談中になればすぐ対応し, 面談中ならばその場で調整します.
- ◆ **Q&A集**も参考にしてください:
<https://www.ueda.info.waseda.ac.jp/faq/>

10

教員や先輩との面談が可能な時間帯は, このスライドを参照してください.
上記のウェブサイトにて, 訪問や面談の最新情報を掲載します.
このウェブサイトからたどれる上田研Q&A集も参考になるかと思えます.