# HARD Spring Worksohp 2021

## Home AI Robot Development Spring Workshop

## 最終回 Speech Recognition

Kosei Demura
Kanazawa Institute of Technology

# HARDのコンセプト

AI Robot

広く浅く全体を俯瞰

Robotics

AI

Computer
Science & Math

Programming
Skill

Computer
Literacy

1. 開発を継続できる能力育成
2. プロと同じテクノロジー
3. 個人でもできる

# 想定する参加者

## 知能ロボットのビギナー

1. ロボットや人工知能に興味のある高校生や学部低回生
2. これから研究室でロボットを使って研究を始める学部生、大学院生
3. RoboCup@Home Leagueに参加したい社会人な

# HARD Spring Workshop 2021

情報工学と数学・力学の礎が大切

# 講師紹介

- 出村公成 (DEMURA, Kosei)
  - 金沢工業大学
    - ロボティクス学科 教授
    - FMT研究所 研究員
    - 夢考房RoboCupプロジェクト指導責任者
    - 夢考房Junior 代表
  - ロボカップジュニア石川ブロック長
  - Folding@home チームKanazawa (ID257261) 代表
    - 世界160位（225,452チーム中）メンバー募集中

- Twitter: demura.net、#HARD2021
- Facebook：出村公成

**ロボット競技会・授業の経験を提供**

# HARDの予定（全4回）

1. 3/6(土): 説明会、AIロボットのつくり方（開発手法）
2. 3/13(土): ロボット制御のつくり方 – 自律移動ロボットの基礎
3. 3/20(土): ロボット視覚のつくり方 – センサ情報処理
   - 特徴量（色、形状）
   - ウェブカメラ、LIDAR
   - 開発：色、形による追跡、距離による追跡、障害物回避
   - 終了後：30分程度質問対応(ブレイクアウトルーム)
4. 3/27(土): ロボット聴覚のつくり方 – 音声認識・合成
   - 音声認識・合成（Web Speech API)
   - 開発：簡単な音声応答システム
   - 今後、開発を続けるために
   - 終了後：ソフトな交流会

**時間：14:00–16:00**
**場所：同じZoom会場（Zoom登録は初回だけ）**

# アジェンダ

1. レクチャー（60分程度）
   - 音声認識
2. ハンズオンの説明(30分)
3. 競技会へ参加しよう！(20分)
   - RoboCup@Home Education 参加方法
   - 10万円でつくる生活支援ロボット
   - 開発を続ける方法
4. アンケート(10分)
5. ソフトな交流会

# @Homeの動画を紹介する

https://youtu.be/RR4tuToroKM

# 音声認識って何だろう？

# 音声認識とは何だろう？

# 音声認識

# 音声認識手法の歴史

| 年代 | 手法 |
| --- | --- |
| 1960～1980年代 | テンプレートマッチング |
| 1980～2000年代 | 統計モデル |
| 2010～ | 一部ニューラルネット |
| 2015～ | 全部ニューラルネット |

# 音声認識

音声の時系列
特徴量 X

⇨

単語列 W

言語モデル　　音響モデル

Xのとき
Wになる確率

その状況で
Wになる確率

Wのとき
Xになる確率

$$P(W|X) \approx P(W)P(X|W)$$

条件付き確率　　ベイズの定理

# 条件付き確率？

# 条件付き確率

2つの対の事象x,yについて，P(x|y)はyが与えられたときにxが起きる確率を**条件付き確率**とよぶ

$$\boxed{P(x \mid y) = \frac{P(x, y)}{P(y)}}$$

*P(x,y): xとyが同時に起きる確率*

# ベイズ定理（ベイズ則）

ある原因X（データ）が得られた時、その結果Yを反映した下での事後確率P(X|Y)を求める方法。
イギリスのトーマス・ベイズが発見。

$$P(X|Y) = \frac{P(Y|X)\ P(X)}{P(Y)}$$

- 条件付き確率の定義より
  - P(X|Y) = P(X,Y)/P(Y)　　(1)
  - P(Y|X) = P(Y,X)/P(X)　　(2)
- P(X,Y)とP(Y,X)は等しいので
  (2)式を(1)式に代入

Thomas Bayes
(1702-1761)

# ベイズ則(最重要)

尤度(ゆうど)　事前確率

posterior

$$P(x|y) = \frac{P(y|x)\ P(x)}{P(y)} = \frac{\text{likelihood} \cdot \text{prior}}{\text{result}}$$

事後確率　　結果

x(原因)→y(結果)
P(x|y): 結果が分かってから原因が起きる確率
事後確率

# 因果的 vs. 診断的　推論（Reasoning）

- *P(open|z)* は 診断的（diagnostic）.
- *P(z|open)* は 因果的(causal)
- 因果的知識は獲得が容易な場合が多い
- ベイズ則は因果的知識を利用して診断的知識を得ることができるところが素晴らしい

頻度を数えろ！

$$P(open \mid z) = \frac{P(z \mid open)P(open)}{P(z)}$$

# おすすめの教科書はこれだ！

ベイズ統計学入門
著者:渡辺洋
出版社 : 福村出版 (1999/9/1)
発売日 : 1999/9/1
言語 : 日本語
単行本 : 249ページ
ISBN-10 : 4571200668
ISBN-13 : 978-4571200663

# おすすめの教科書はこれだ！

確率ロボティクスのバイブル

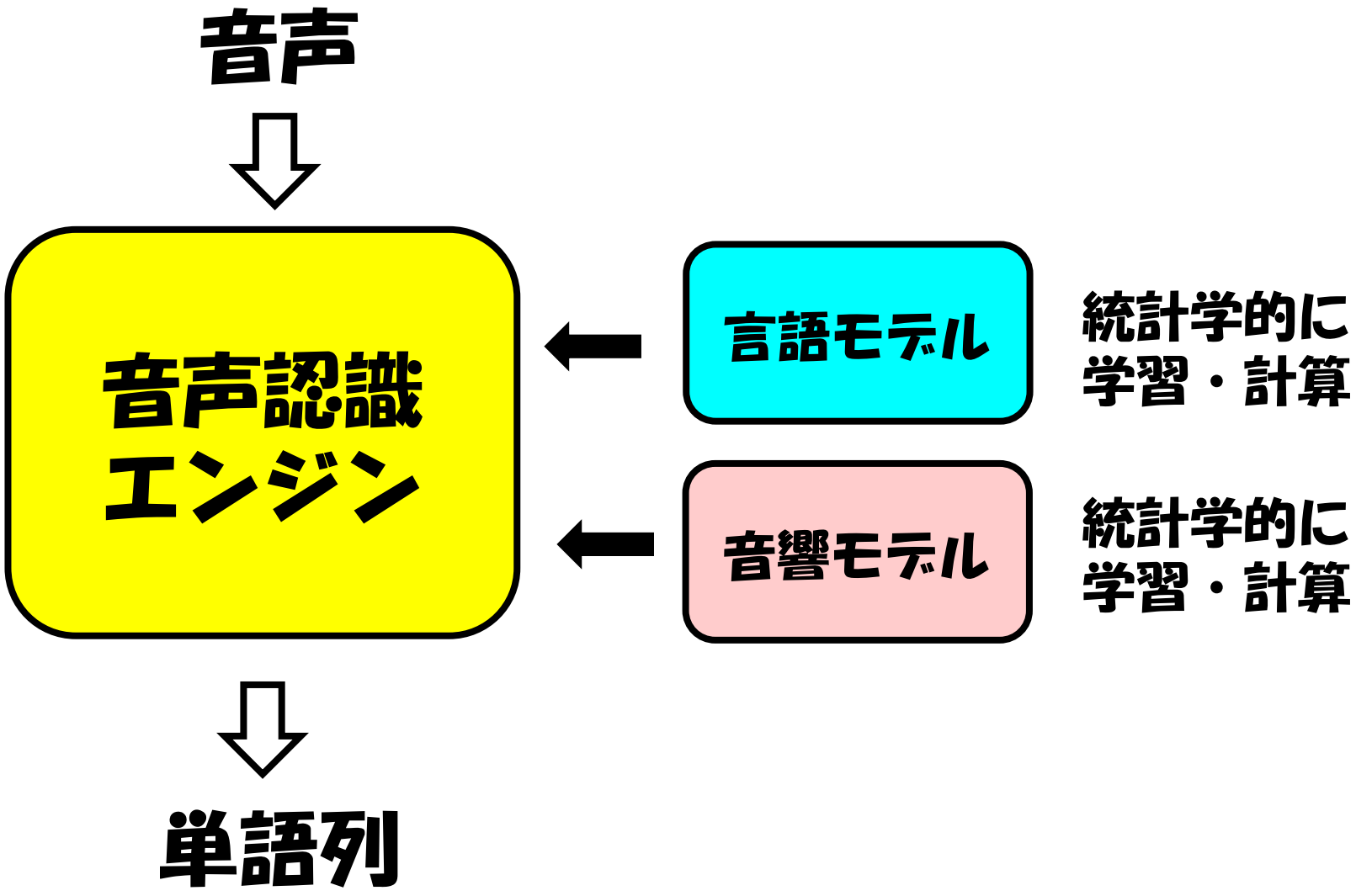
- 書名：Probabilistic ROBOTICS
- 著者：Thrun, Burgard, Fox , 2005
- 発行：The MIT Press
- 価格：9160円

9160円　　9500円

詳解　確率ロボティクス
Pythonによる基礎アルゴリズムの実装, 上田隆一 (著)

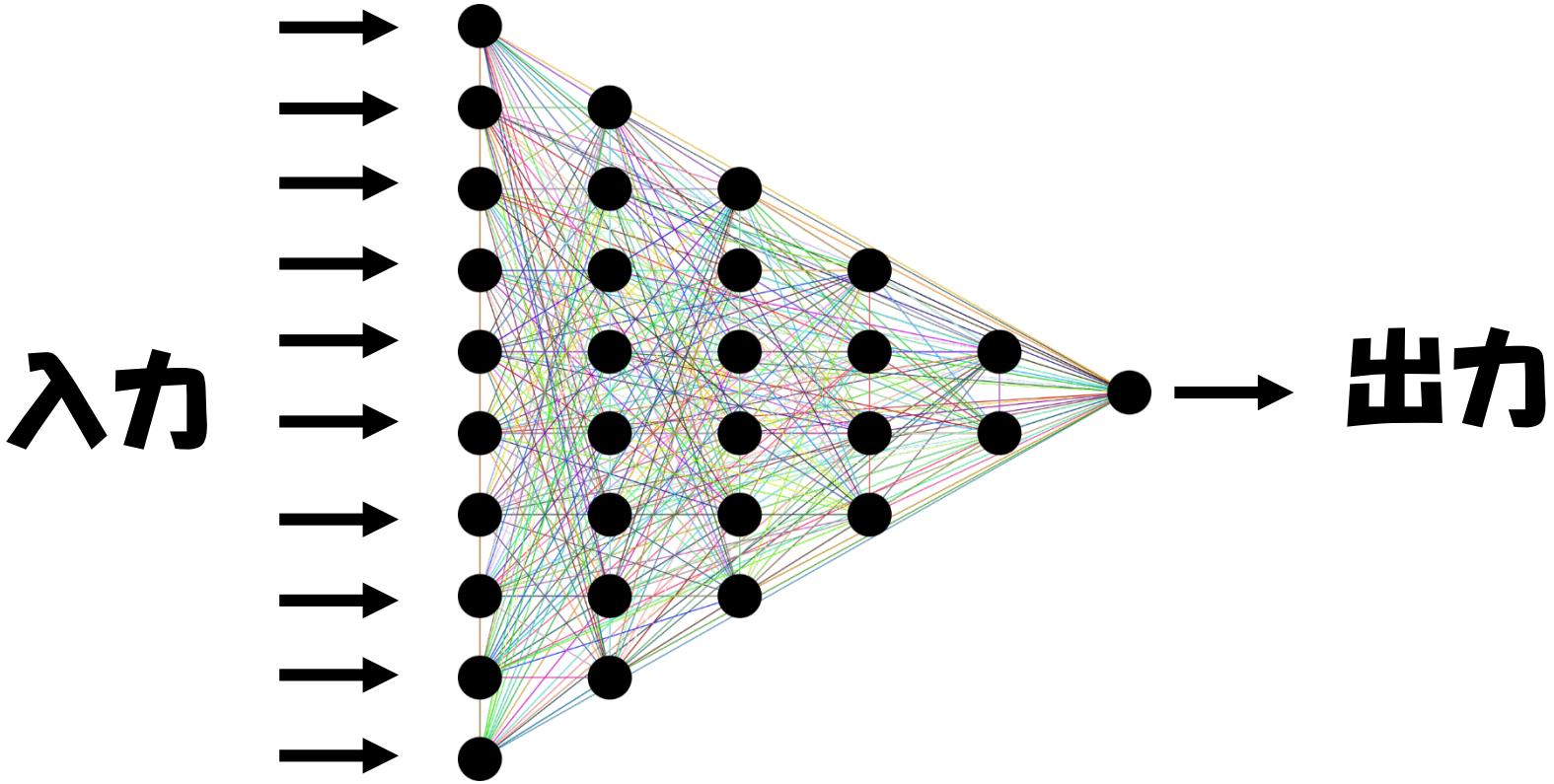
4290円

# 統計モデルの音声認識システム



# ニューラルネットワーク？

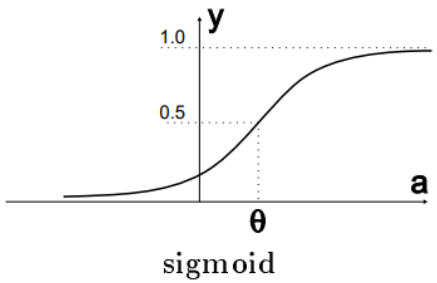
# ニューラルネットワーク

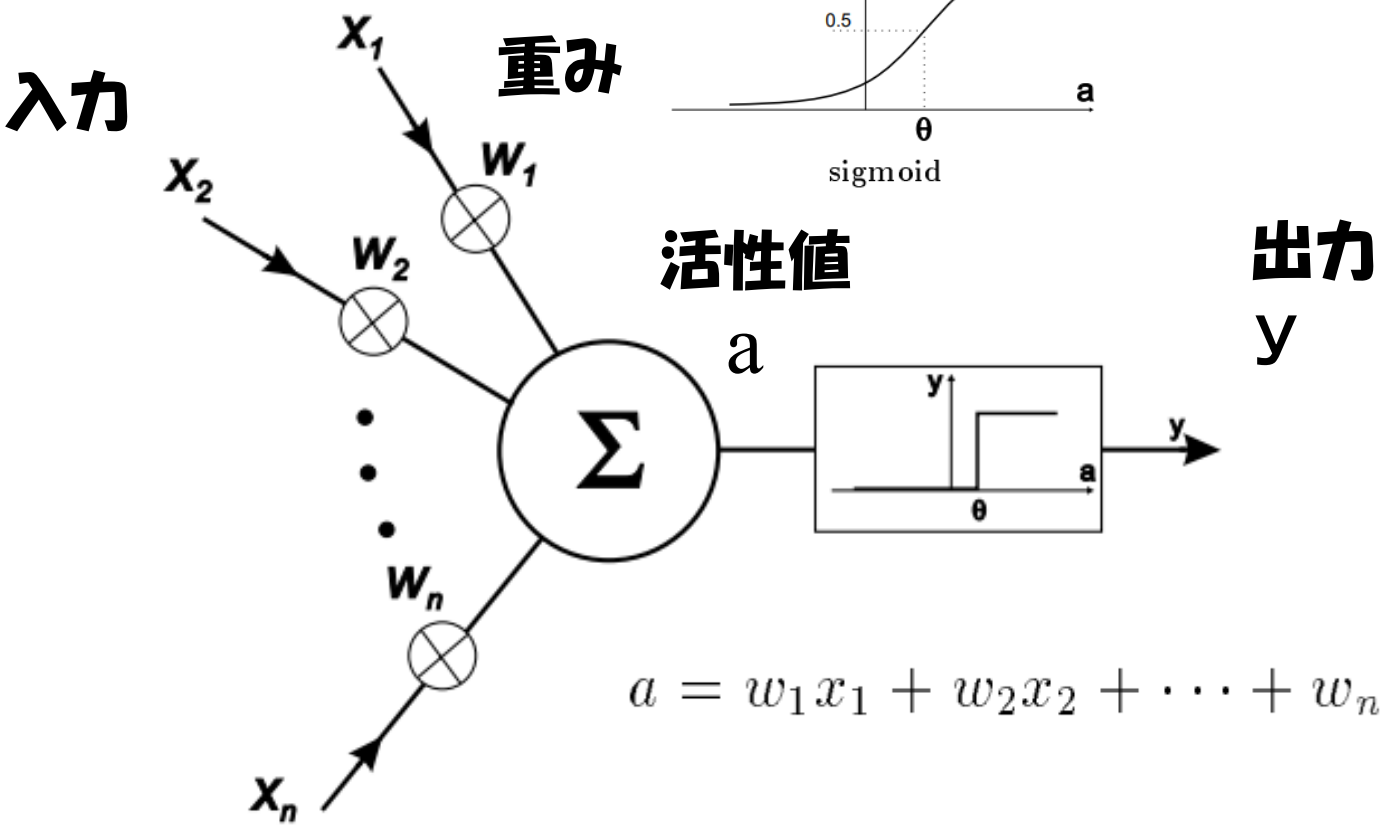



生物の神経回路網を工学的に実現したもの

# モデル

$$f(x) = \frac{1}{1+e^{-ax}} \quad (a > 0)$$




$$a = w_1 x_1 + w_2 x_2 + \cdots + w_n x_n$$

$$y = \begin{cases} 1 & \text{if} \quad a \geq \theta \\ 0 & \text{if} \quad a < \theta \end{cases}$$

$$\sum_{k=1}^{n} k = 1 + 2 + 3 + \cdots + n$$

# 教師あり学習

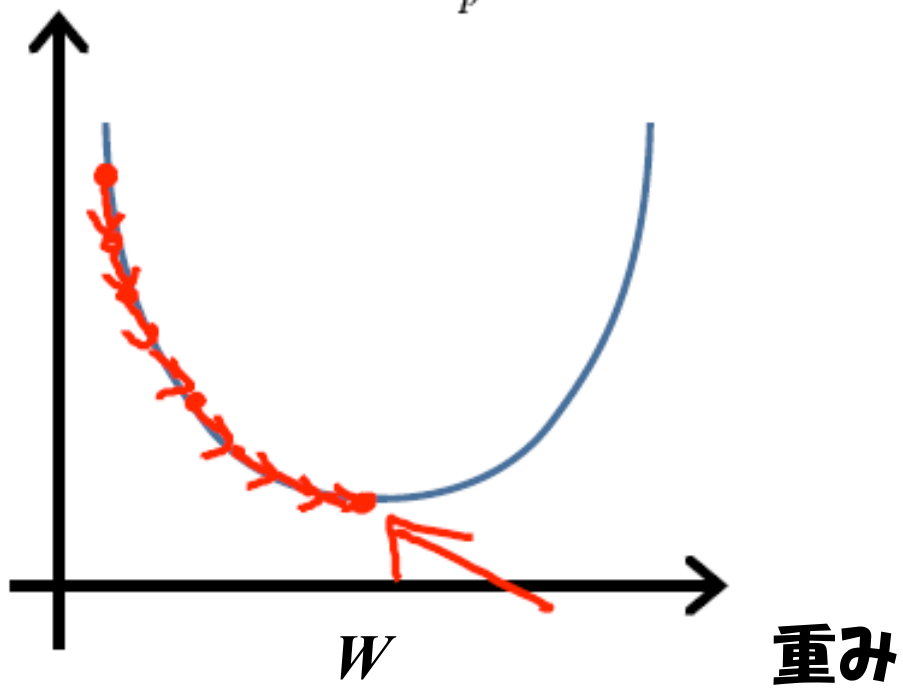
入力p時

$$E_p = \frac{1}{2}(t - y)^2$$

誤差関数

$$E = \sum_p E_p$$

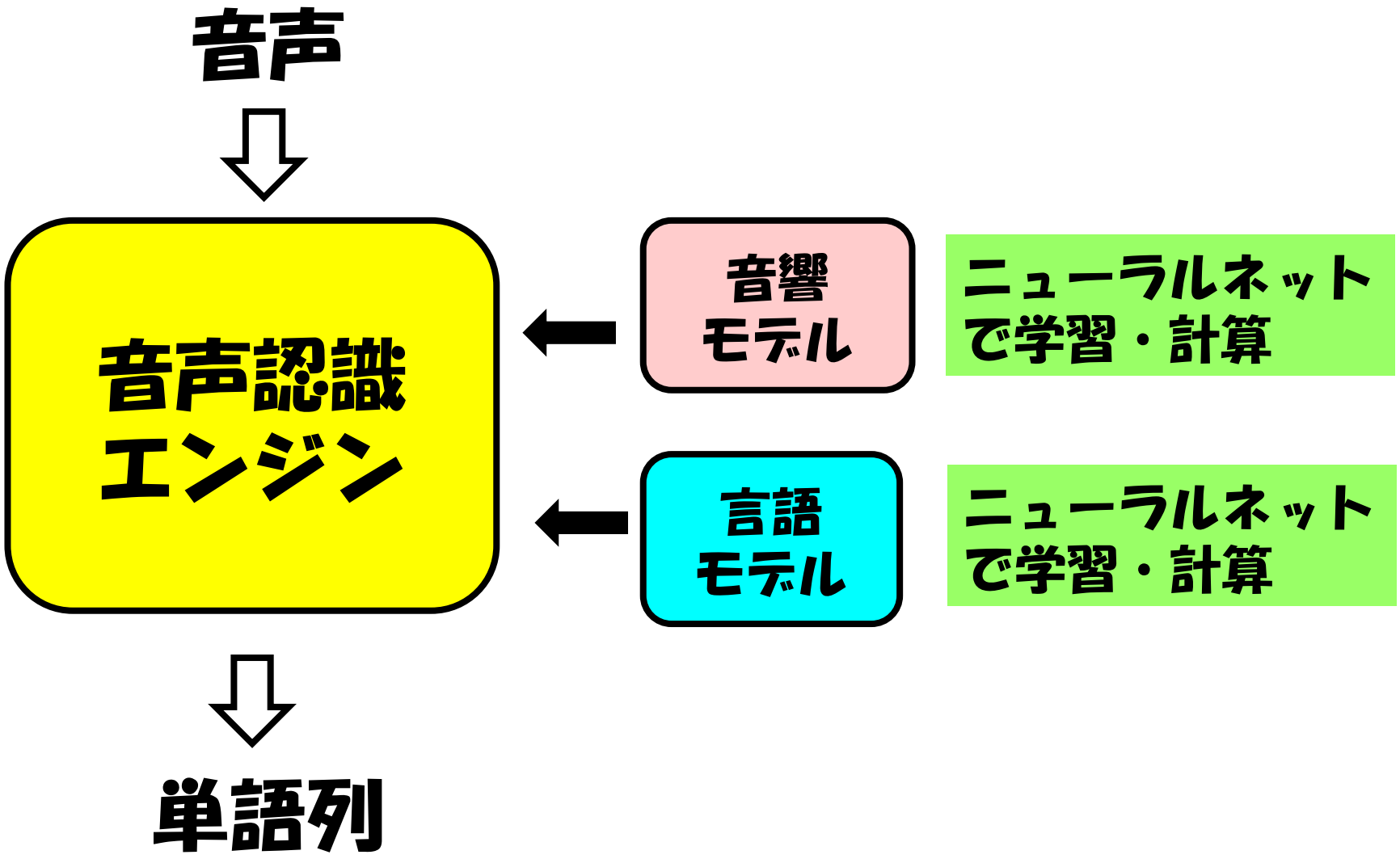
関数の傾きを調べて
重みを繰り返し修正




誤差関数E ＝（教師値t － ネットの出力y）の二乗和

誤差関数を最小にする重みを見つけるのが学習

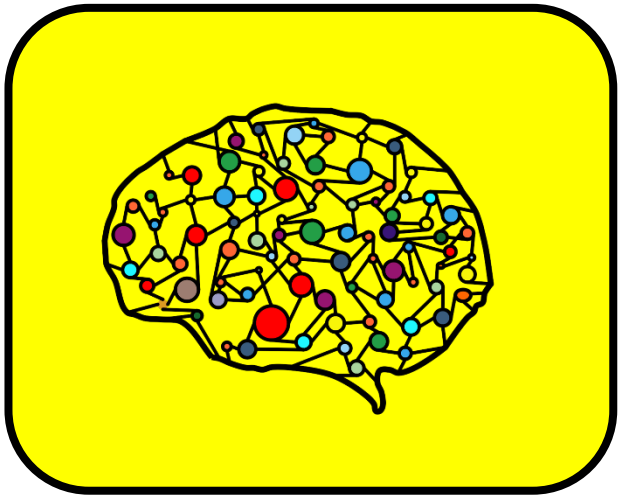
# 初期ニューラルネットの音声認識システム



# 最新ニューラルネットの音声認識システム

音声

⇩



⇦

全ての計算を
ニューラルネット
で学習・計算

End-to-Endシステム

⇩

単語列

# これで何が起きたか？

人間と同等の認識精度を達成！

# アジェンダ

1. レクチャー（60分程度）
   - 音声認識
2. ハンズオンの説明(30分)
3. 競技会へ参加しよう！(20分)
   - RoboCup@Home Education 参加方法
   - 10万円でつくる生活支援ロボット
   - 開発を続ける方法
4. アンケート(10分)
5. ソフトな交流会

# アジェンダ

1. レクチャー（60分程度）
   - 音声認識
2. ハンズオンの説明(30分)
3. 競技会へ参加しよう！(20分)
   - RoboCup@Home Education 参加方法
   - 10万円でつくる生活支援ロボット
   - 開発を続ける方法
4. アンケート(10分)
5. ソフトな交流会

# RoboCup@Home Eduに参加するには？

# Jeffrey先生

# RoboCup@Home Edu Founder

Japan Open @Home Edu実行委員長

# 萩原先生　立命館大学　講師

# アジェンダ

1. レクチャー（60分程度）
   - 音声認識
2. ハンズオンの説明(30分)
3. 競技会へ参加しよう！(20分)
   - RoboCup@Home Education 参加方法
   - 10万円でつくる生活支援ロボット
   - 開発を続ける方法
4. アンケート(10分)
5. ソフトな交流会

# 音声認識ソフトウェア紹介

# 音声認識エンジン

- オンライン
  - 無料
    - Web Speech API
    - Google Speech Recognition(テスト用)
    - DeepSpeech
  - 有料
    - Google Cloud Speech API
    - Microsoft Bing, IBM Speech to Text, Wit.ai
- オフライン
  - 無料
    - Julius
    - Kaldi
    - Pocket sphinx
    - Flashlight (facebook)
    - DeepSpeech (Baiduのクラウド)
  - 有料
    - ?

# 開発を続けるには

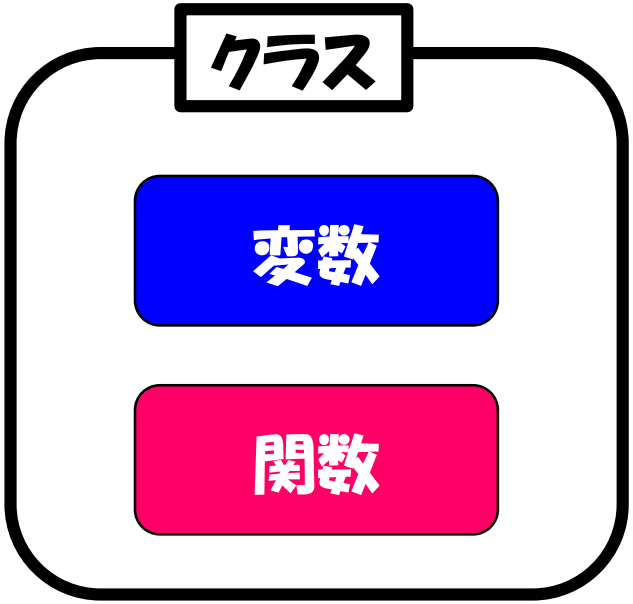
# C言語のプログラミング

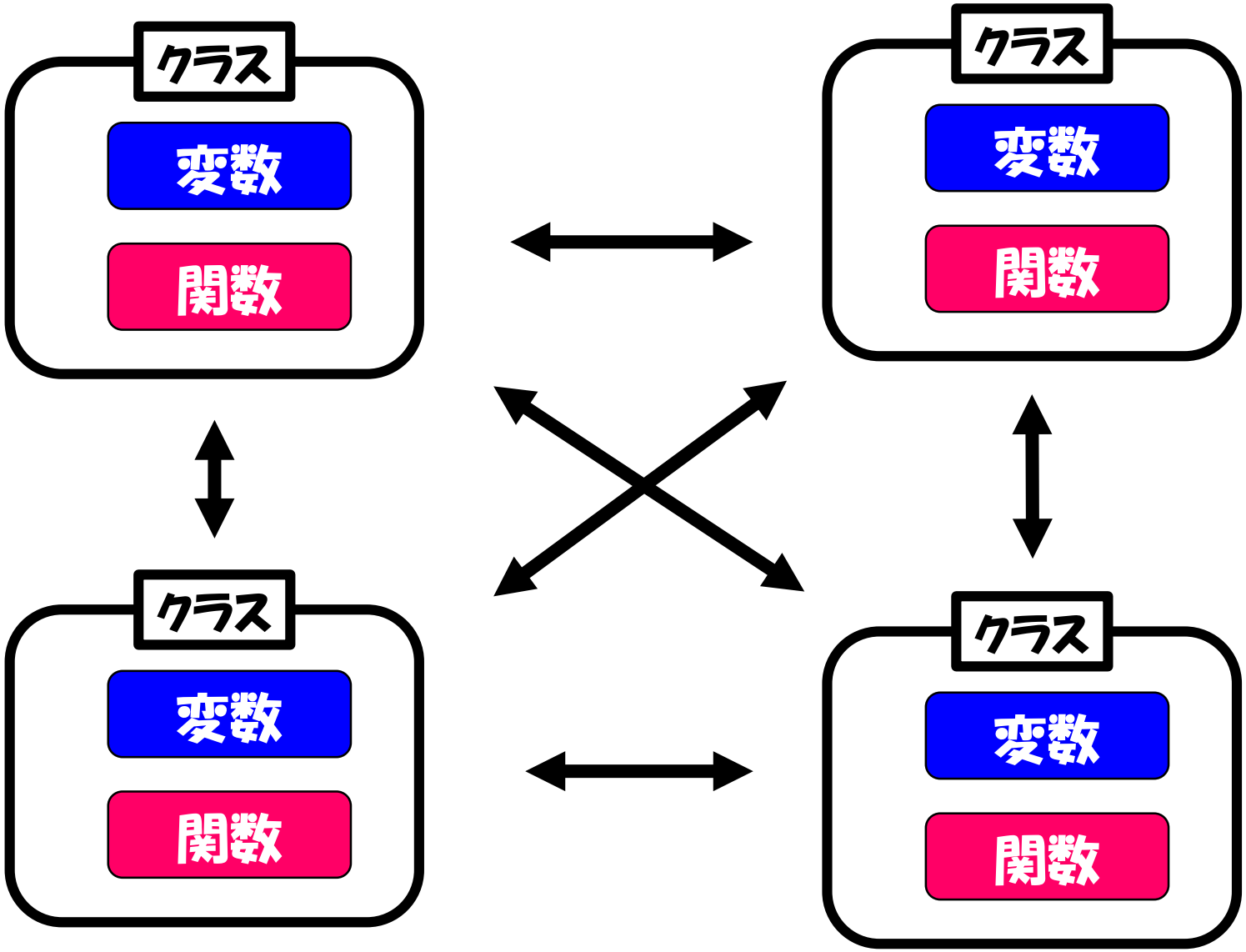
# 関数を最小単位としてプログラミング

# オブジェクト指向言語のプログラミング

## C++、Python、JAVA

# クラスを最小単位としてプログラミング

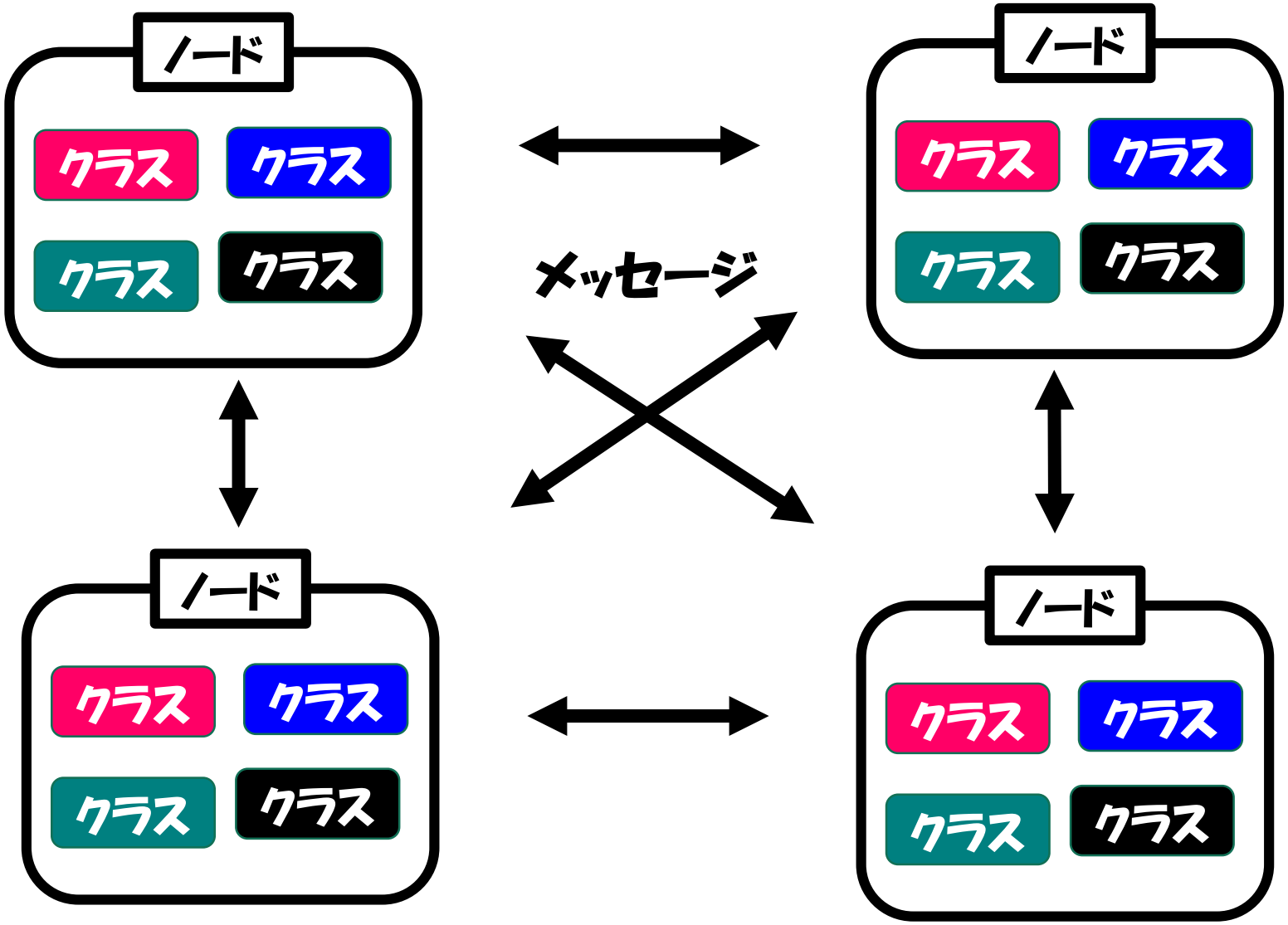



クラス
変数
関数
クラス
変数
関数
クラス
変数
関数
クラス
変数
関数

# 知能ロボットのプログラミング

# ROS

# ノードを最小単位としてプログラミング



# 問題の解決にノードをどう組みわせるか

既存ノードを使うだけでは力がつかない。

できるだけ自作ノードを実装しよう！

ムズ！

ソースを読むことからはじめよう！

Python、C++

# 独学のすゝめ

# 競技会参加のすゝめ

モチベーション

ネットワーキング

アンケート　10分

チャットでURLを送付

# 10万円で作る!?

# Home AI ロボット

これがおススメのハードウェアだ！

# おすすめマイクはこれだ！

## ガンマイク

RODE ロード VideoMic GO,

## マイクアレー

ReSpeaker Mic Array v2.0, Seeed,

# おすすめスピーカはこれだ！

- 40W
- アナログ入力
- 小型 20×6×8[cm]
- 充電式
- Bluetoothあり

Tronsmart Bluetooth5.0スピーカー 40W

# 今、おすすめスピーカシステムこれだ！

- **100W(50W×2)**
- アナログ入力あり
- 小型 8×7×4[cm]
- **DC 9～24V必要**
- Bluetoothあり

ELEGIAN　Bluetooth 5.0
ステレオ スピーカー パワーアンプ

FOSTEX スピーカーユニット
P800K (8cm)
P1000K(10cm)

# おすすめカメラはこれだ！

ASUS ROG Eye

ロジクール ウェブカメラ C920n

# おすすめLIDARはこれだ！

シミュレータと同じ

RPLiDAR A1M8

RPLiDAR A2M8

# おすすめのRGB-Dセンサはこれだ！

Intel Realsense Depth Camera D435i

# おすすめの台車はこれだ！

ルンバ606

- 600～800シリーズ（シリアルコネクタあり）
- 掃除できる（自分で掃除ロボのソフト開発）
- 通信ケーブルを自作、購入
- 最新シリーズ692は対応不明

Create2

- 通信ケーブル付き
- 掃除はできない

# おすすめのアームはこれだ！

RT CRANE+V2 ROS対応 アームロボット
RTロボットショップ

# おすすめのアームは自作だ！

サーボモータ

ブラケット

**ROBO-ONE参加者のページ参考**

# おすすめの外装は自作だ！

・3Dプリンタ
・発砲スチロール
・樹脂、布、紙

Fig.1 Happy Mini 2016

Honorin

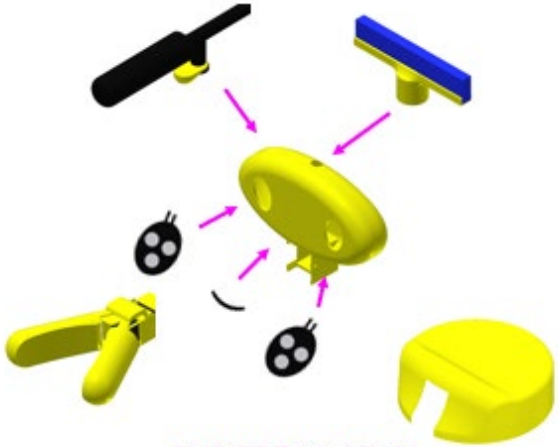


Fig.2 3D Printed Parts

Happy mini

**とても重要！**

# How much ?

| 必要必要性 | | | |
|---|---|---|---|
| 必須 | PC | Raspberry Pi4 | 7,000 |
| 必須 | 台車 | ルンバ606 | 30,000 |
| 余裕があれば | アーム | CRANE+ V2 | 72,600 |
| | LIDAR | RPLiDAR A1M8 | 10,000 |
| | | | |
| | カメラ | Webカメラ | 10,000 |
| 欲しい | マイク | | |
| 欲しい | スピーカ | | |

# 質　問

おわり