

ことばを理解すること なぜコンピュータには難しいのか？

宮尾祐介

国立情報学研究所

コンテンツ科学研究系

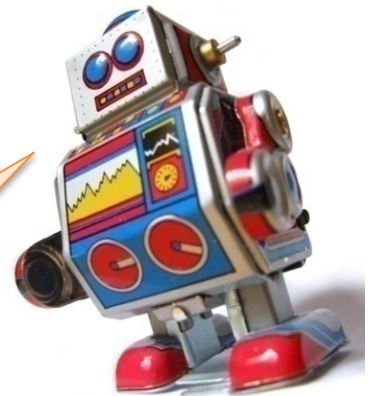
自然言語処理

- 人間のことば(自然言語)を理解するコンピュータを作ることを目指す学問
- ことばによる情報交換、コミュニケーションを助ける
 - かな漢字変換
 - 検索
 - 自動翻訳
 - 質問応答



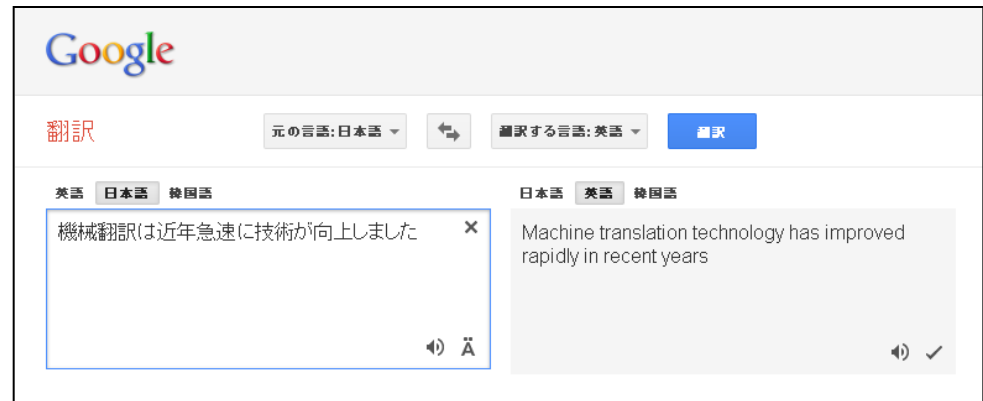
おいしいイタリアン
知らない？

いいところ知ってますよ。
高級店とカジュアル店の
どっちがいいですか？

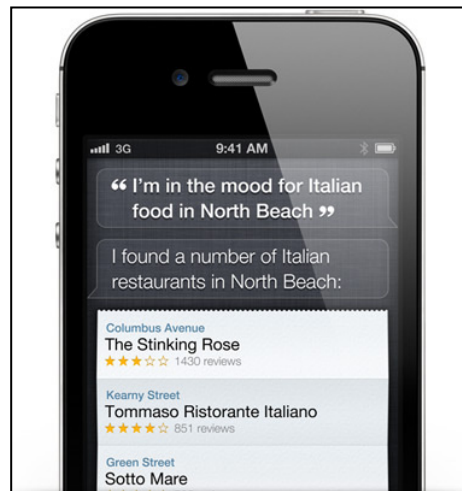


自然言語処理の応用

かな漢字変換



自動翻訳



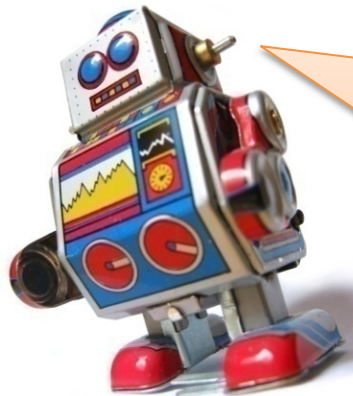
質問応答

コンピュータがことばを理解しはじめている？

「ことばを理解する」とは？

- 自然言語の入力に対して、「意味のある」反応をする
- 「意味のある」反応とは？
 - 会話ボット
 - 検索
 - 質問応答

おいしいイタリアン
知らない？



いいところ知ってますよ。
高級店とカジュアル店の
どっちがいいですか？

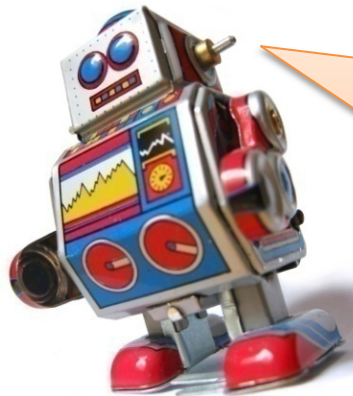
入力

出力

「ことばを理解する」とは？

- 自然言語の入力に対して、「意味のある」反応をする
- 「意味のある」反応とは？
 - 会話ボット
 - 検索
 - 質問応答

おいしいイタリアン
知らない？



AKB総選挙について
教えてくださいませんか？

入力

出力

検索・質問応答のしくみ

- 検索・質問応答のすべき反応
＝ユーザがほしい情報を探す
- だれかがどこかで書いている文書を見つける
→ キーワードが含まれる文書を探す

文書1

昨日食べた魚
は腐ってた
みたい

文書2

おいしいスイカ
をたくさん
食べた

スイカ

文書3

みかんを一箱
もらったけど
全部食べた

文書4

実家のスイカ
をたくさん
もらった

文書5

実家で作ってる
みかんは全部
おいしいよ

昨日	1
食べた	1, 2, 3
魚	1
腐ってた	1
おいしい	2, 5
スイカ	2, 4
たくさん	2, 4
みかん	3, 5
一箱	3
もらった	3, 4
全部	3, 5
実家	4, 5
作ってる	5

「ことばを理解している」 と言っているのか？

- いいんです！
 - 目的とする「意味のある」反応ができていれば
- でも、やっぱりもうちょっと「人間っぽく」理解させたい
 1. もっと難しい「意味ある反応」にチャレンジする
 2. 「人間っぽい理解」とは何？

ロボットは東大に入れるか？

- 人工知能で東大入試を突破することを目指すプロジェクト
- 自然言語の理解が重要な役割を果たす
- 「意味のある」反応＝問題に正しく解答する！

問題文

第 4 問

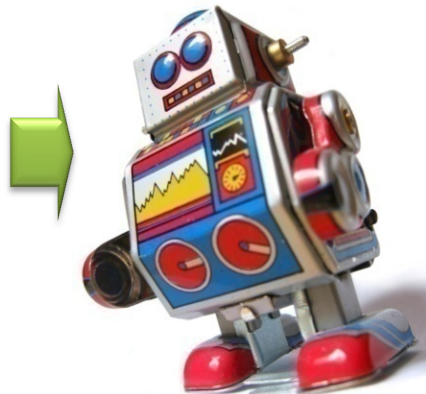
O を原点とする座標平面上の曲線

$$C: y = \frac{1}{2}x + \sqrt{\frac{1}{4}x^2 + 2}$$

と、その上の相異なる 2 点 $P_1(x_1, y_1)$, $P_2(x_2, y_2)$ を考える。

(1) $P_i (i=1, 2)$ を通る x 軸に平行な直線と、直線 $y=x$ との交点を、それぞれ $H_i (i=1, 2)$ とする。このとき $\triangle OP_1H_1$ と $\triangle OP_2H_2$ の面積は等しいことを示せ。

(2) $x_1 < x_2$ とする。このとき C の $x_1 \leq x \leq x_2$ の範囲にある部分と、線分 P_1O , P_2O とで囲まれる図形の面積を、 y_1, y_2 を用いて表せ。



解答

(1) $P_i (x_i, y_i)$ を通る x 軸に平行な直線と、 $y=x$ との交点は $H_i(y_i, y_i)$ であるから、

$$\triangle OP_iH_i = \frac{1}{2}|(y_i - x_i)y_i|$$

ここで、

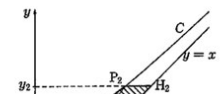
$$\begin{aligned} (y_i - x_i)y_i &= \left(\frac{1}{2}x_i + \sqrt{\frac{1}{4}x_i^2 + 2} - x_i\right)\left(\frac{1}{2}x_i + \sqrt{\frac{1}{4}x_i^2 + 2}\right) \\ &= \frac{1}{4}x_i^2 + 2 - \frac{1}{4}x_i^2 \\ &= 2 \cdots \textcircled{1} \end{aligned}$$

よって、 $\triangle OP_iH_i = \frac{1}{2} \cdot 2 = 1 (i=1, 2)$ であるから、 $\triangle OP_1H_1$ と $\triangle OP_2H_2$ の面積は等しい。

(2) $\frac{1}{2}x + \sqrt{\frac{1}{4}x^2 + 2} > \frac{1}{2}x + \frac{1}{2}|x| \geq 0$

よって、 C は $y > 0$ を満たす領域にあるから、①より、曲線 C の方程式は

ら、①より、曲線 C の方程式は

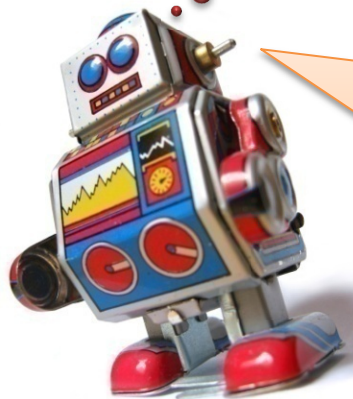


人間は何をやっている？

- 入力：自然言語
- 出力：期待する反応（欲しい情報、解答、...）
- 入力と出力の間に何かが起きている
＝言葉の理解

おいしいイタリアン
知らない？

入力



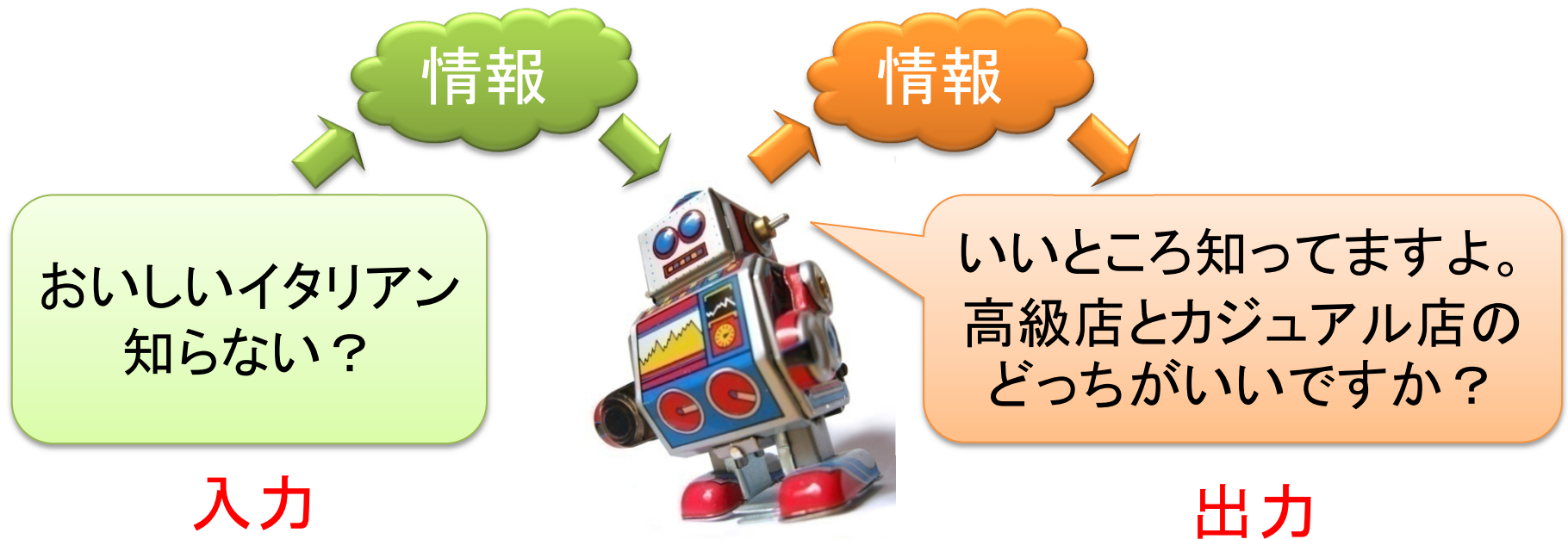
理解

いいところ知ってますよ。
高級店とカジュアル店の
どっちがいいですか？

出力

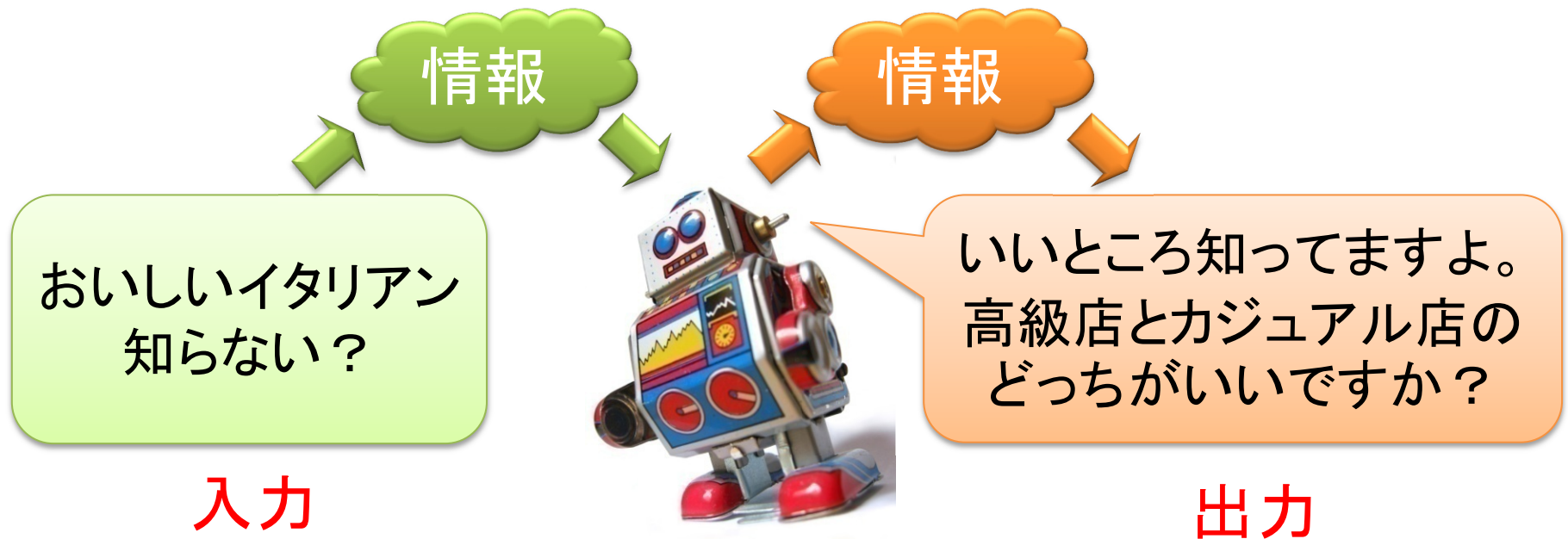
情報のやりとり

- 人間は言葉で「情報」をやりとりしている
 - 情報を提供し、情報を獲得する
- 言葉が表す情報(=「意味」)を分かってほしい
→ 意味理解



「意味」とは何か？

- なぜ言葉を文字列そのままに処理するのではなく、「意味」を考える必要があるのか？



「意味」とは何か？

今日はあくまで
「ことばが表す情報」が対象

- 「意味」(大辞林)

1. 言葉・記号などで表現され、また理解される一定の内容。

2. ある表現・作品・行為にこめられた内容・意図・理由・目的・気持ちなど。

3. 物事がある脈絡の中でもつ価値。重要性。意義。

- どうやってコンピュータで実現すればよいのか？

- たぶん人間の頭の中にあるが、どういうものか観察できない

→ どういうものか分からないので、**その性質から迫る**

自然言語処理における 「意味」とは何か？

- 異なる文字列が同じ「意味」を表す
 - まる子がケーキを食べた。
 - まる子がケヤキをくべた。
 - モンブランをまる子に食べられた。
 - 同じ文字列が異なる「意味」を表す
 - かわいい瞳の大きな女の子を見た。
 - つまり、意味が「同じ(同値性)」「異なる(差異)」という直感を再現したい
- 「同じ」「異なる」をコンピュータで計算できるように、**意味の表現方法・計算方法**を設計する

かわいい瞳
の大きな女

...の子ども



文字列から意味へ

犬に風邪薬を飲ませると貧血状態に陥ります。

どういうときに「意味が同じ」と
言えるのか？

ビーグルが風邪薬を食べたら病気になる！



メイちゃん(ビーグル♀)

意味の計算

犬に風邪薬を飲ませると
貧血状態に陥ります。

どうやってつなぐ？

意味表現

ビーグルが風邪薬を食べたら病気になる

意味の2つの側面

- **構成的意味**：文の中で単語が組み合わされて表現される意味
 - － ビーグルが風邪薬を飲んだ。
 - － ビーグルに風邪薬を飲ませた。
 - － それがビーグルに飲ませた風邪薬だ。
- **語彙的意味**：文とは独立に、単語がもともと持っている意味
 - － 犬、ビーグル、柴犬
 - － 顔、目、鼻
 - － あげた、もらった

つまり...

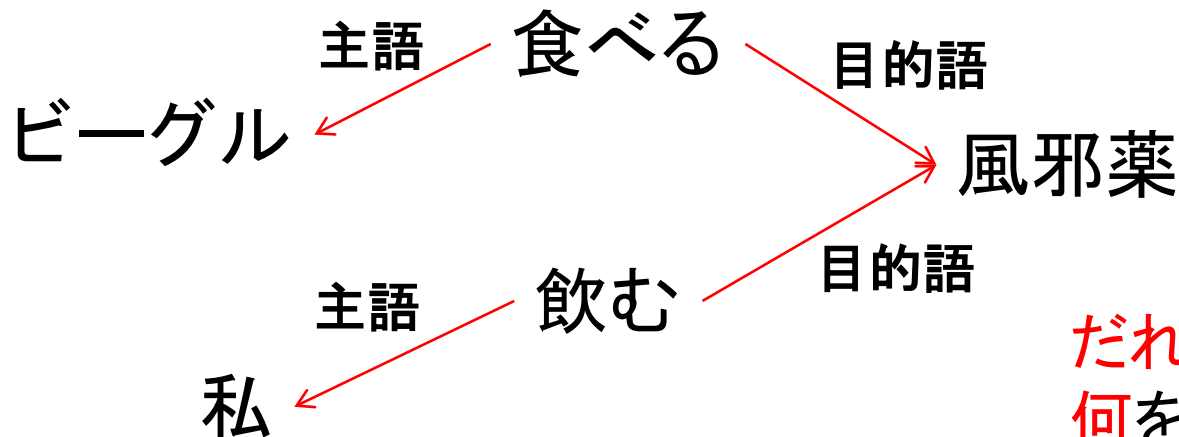
- 言葉が表す意味とは、
 - 単語を並べることによって表される意味と
 - 単語がもともと持っている意味
- の組み合わせ
- この二つを表現し、意味が「同じ」「違う」を計算する仕組みを考えればよい
 - 意味のデータ構造とアルゴリズム

構成的意味

- 単語を並べることによって表される意味
- ビーグルが食べた風邪薬を私も飲んだ。
 - だれが食べた？
 - 何を食べた？
 - だれが飲んだ？
 - 何を飲んだ？
- 私は風邪薬を飲んだ。
- 単語と単語のつながり
 - 主語、目的語、...

構成的意味の表現方法

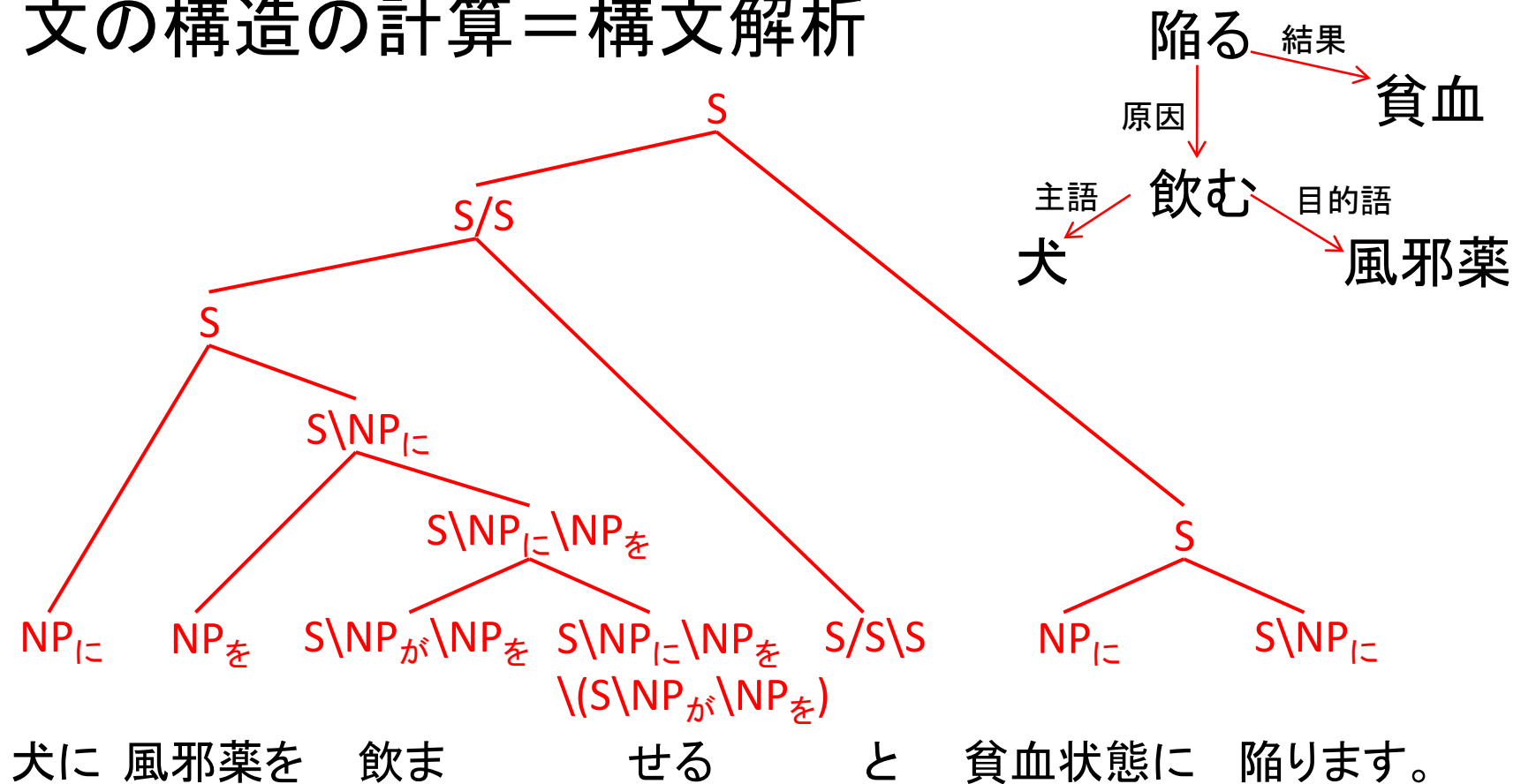
- グラフ構造で表す
 - 点＝単語、辺＝単語間のつながり
 - ビーグルが食べた風邪薬を私も飲んだ。



だれが食べた？
何を食べた？
だれが飲んだ？
何を飲んだ？

構文解析

- 構成的意味は文の構造が分かれば計算できる
- 文の構造の計算＝構文解析



構文解析の研究

- Enju: 英語の構文解析器
- 90%の精度で意味表現(構成的意味)を計算
- 中国語・日本語も研究中

Enju+SCT online demo

Enter sentences, and you will see semantic representations.

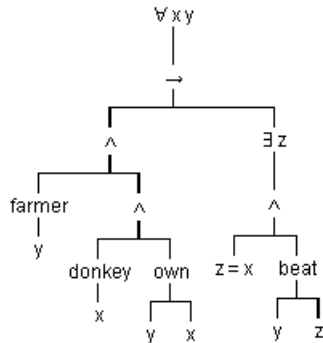
Every farmer that owns a donkey beats it.

Parse

Clear

Click to change the view

$\forall x y ((\text{farmer}(y) \wedge \text{donkey}(x) \wedge \text{own}(y, x)) \rightarrow \exists z (z = x \wedge \text{beat}(y, z)))$



意味の計算

言葉

犬に風邪薬を飲ませると
貧血状態に陥ります。

構文解析

Enju+SCT online demo

Enter sentences, and you will see semantic representations.

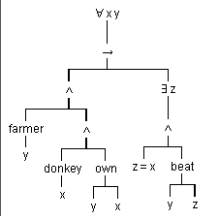
Every farmer that owns a donkey beats it.

Parse

Clear

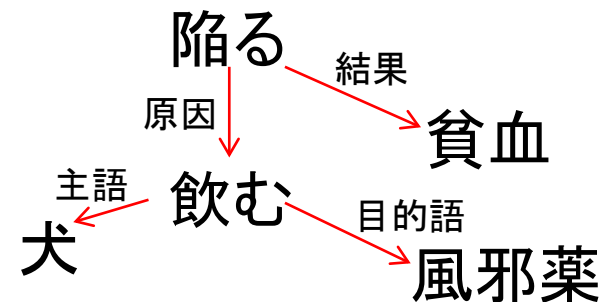
Click to change the view

$\forall x y ((\text{farmer}(y) \wedge \text{donkey}(x) \wedge \text{own}(y, x)) \rightarrow \exists z (z = x \wedge \text{beat}(y, z)))$



構成的意味を計算

意味表現



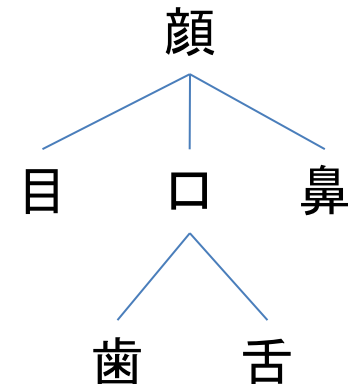
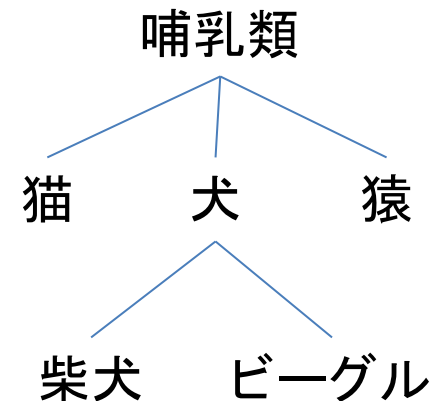
ビーグルが風邪薬を食べたら病気になる

語彙的意味

- 単語がもともと持っている意味
- いろいろな種類の意味的關係
 - － 同義・反義關係
 - 犬、イヌ、わんちゃん
 - － 上位・下位關係
 - 犬、ビーグル、柴犬
 - － 全体・部分關係
 - 顔、目、鼻

語彙的意味を表す方法

- 意味ネットワーク: 単語の意味的關係をグラフ構造で表す
 - 点: 単語
 - 辺: 意味的關係
- いろいろな種類の関係
 - 同義・反義関係
 - 上位・下位関係
 - 全体・部分関係



同義関係

- 同義: (ほとんど) 全ての文脈で置き換え可能な単語

– 車、自動車

– 二酸化炭素、炭酸ガス、CO₂

車 ——— 自動車

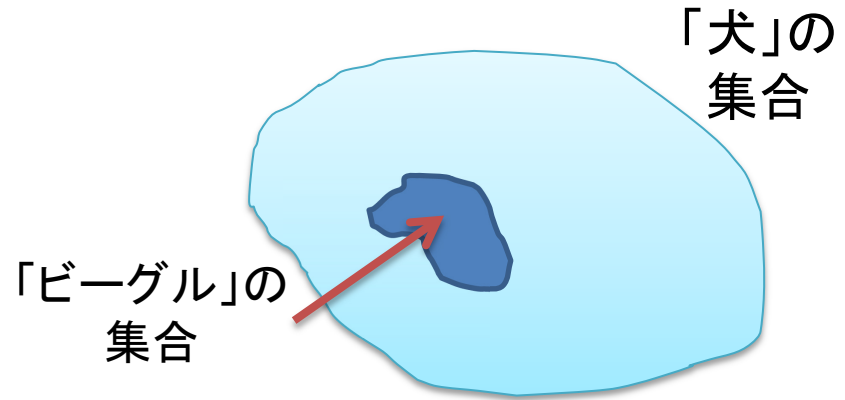
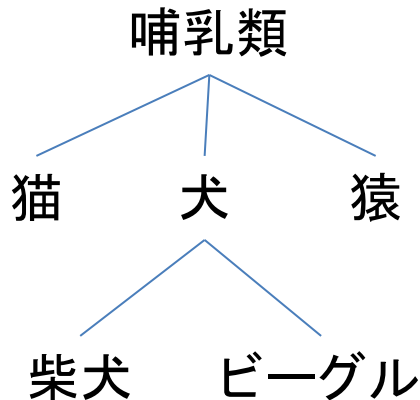
二酸化炭素

炭酸ガス ——— CO₂

冬休みに **x** で草津まで行く。
駅前道はたくさん **x** が走っている。
今度 **x** を修理に出すつもりだ。
昨夜家の前に赤い **x** がとまっていた。

上位・下位関係

- 単語 A の全ての性質を単語 B が持っているなら、A は B の上位語
 - 哺乳類 > 犬、猫
 - 家具 > テーブル、ソファ
- 「B は A の一種である」「B is a A」
- A が指す集合が B が指す集合より大きい



全体・部分関係

- 物理的に「含む・含まれる」関係
 - 顔 > 目, 鼻, 口
- 上位・下位関係と似て非なる関係
 - 白いビーグル → 白い動物
 - 赤い鼻 ≠ 赤い顔
 - 鼻のキズ → 顔のキズ
- 置き換え可能性に関して、上位・下位関係とは異なる性質

動詞の意味関係

- 同義・反義、上位・下位、全体・部分関係は同様に定義できる
 - 勉強する $\equiv$ 学ぶ
- ただし、**項の対応関係**を考える必要がある

X が Y に Z をあげる



X が Y に Z をもらう

あげる(X, Y, Z) = もらう(Y, X, Z)

- 動詞特有の関係もある
 - 含意関係: A が成り立つなら、必ず B も成り立っている
 - X が Y に陥る $\rightarrow$ X が Y になる
 - X を後悔する $\rightarrow$ X が起きた

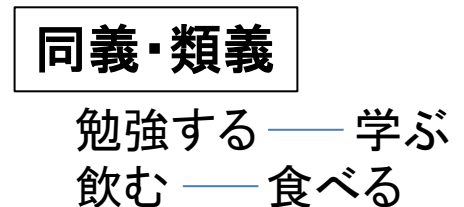
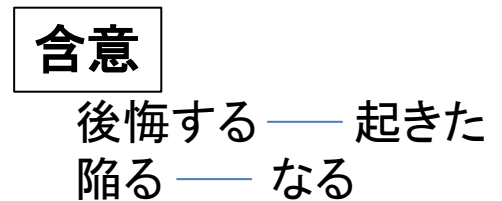
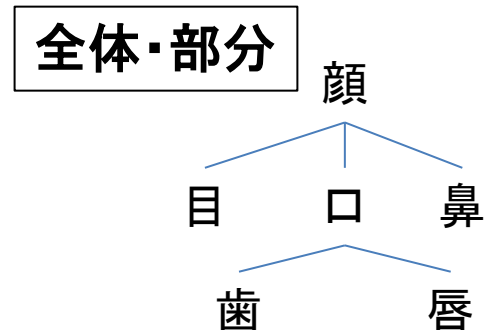
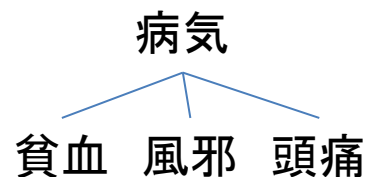
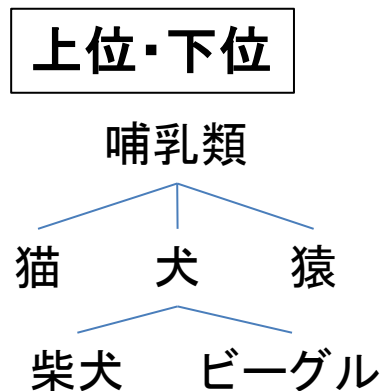
フレーズの意味的關係

- 単語だけでなく、フレーズの間にも意味的關係が考えられる
 - 強い雨が降る = 大雨になる（同義關係）
 - いびきをかく → 寝ている（含意關係）
 - ノーベル文学賞を受賞する
→ 作家である（含意關係）

シソーラス、オントロジー

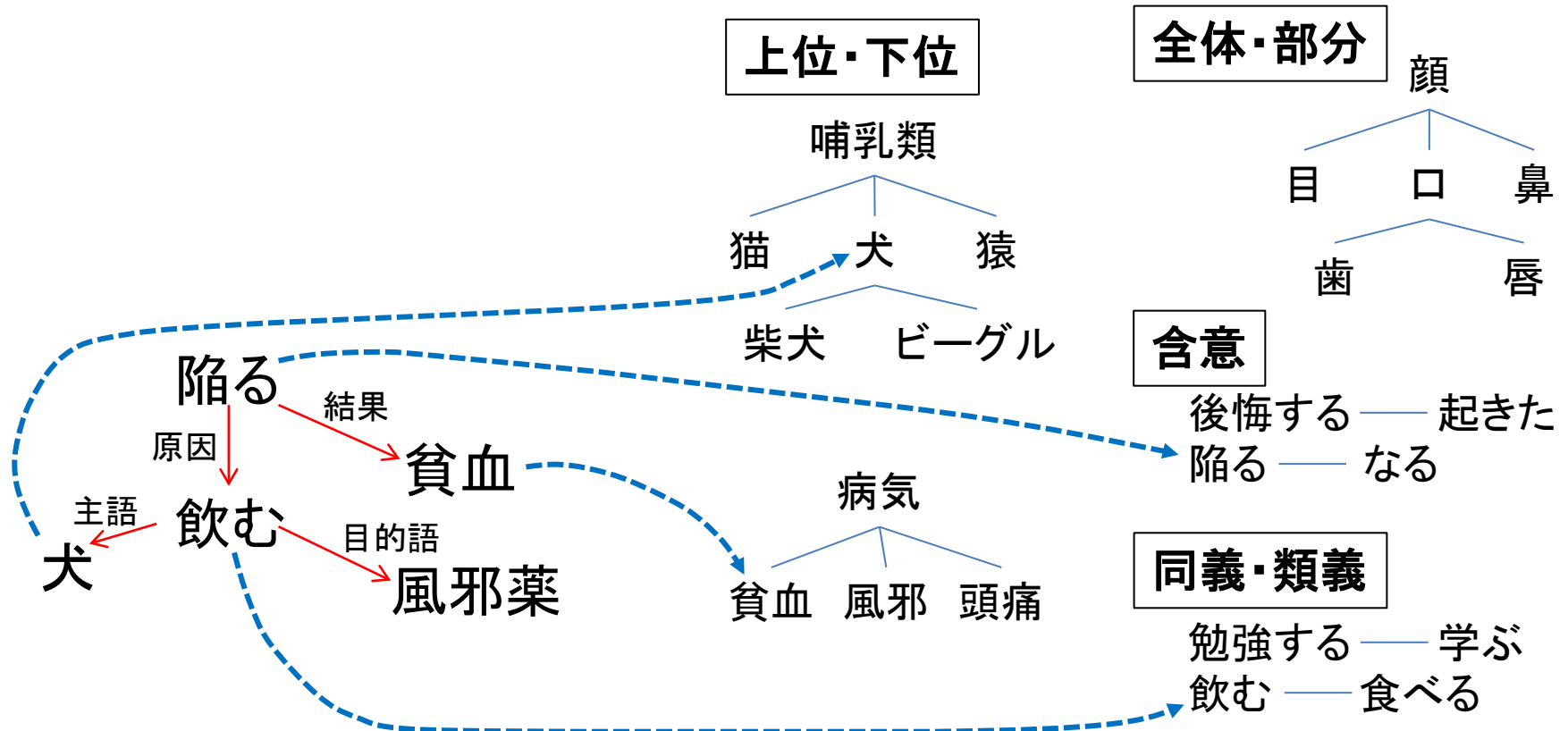
- 同義・反義関係、上位・下位関係などを収録した辞書
- 自然言語処理では欠かせないデータ

- WordNet
- 分類語彙表
- 日本語語彙体系



語彙的意味の計算

- 文中の単語をシソーラス・オントロジーにひもづける
→ 単語の置き換え可能性を表す



意味の計算

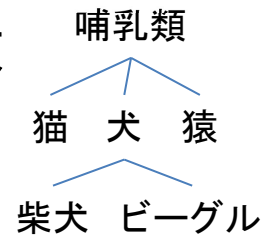
言葉

犬に風邪薬を飲ませると
貧血状態に陥ります。

シソーラス・オントロジー

語彙の意味を計算

上位・下位
全体・部分



同義・類義

飲む — 食べる

含意

陥る — なる

構文解析

Enju+SCT online demo

Enter sentences, and you will see semantic representations.

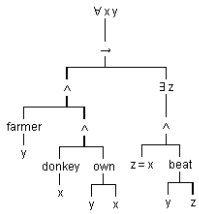
Every farmer that owns a donkey beats it.

Parse

Clear

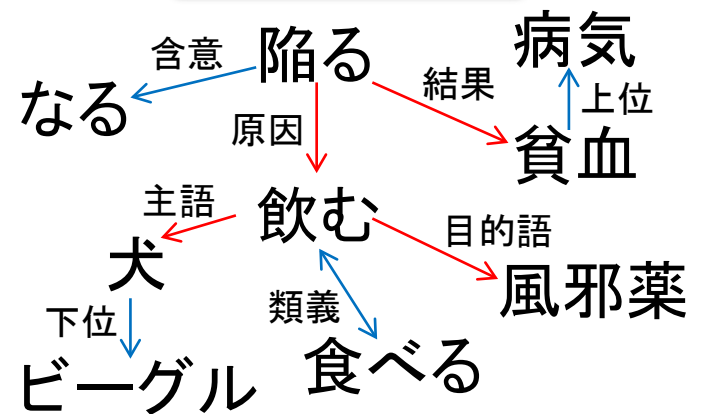
Click to change the view

$\forall x y ((\text{farmer}(y) \wedge \text{donkey}(x) \wedge \text{own}(y, x)) \rightarrow \exists z (z = x \wedge \text{beat}(y, z)))$



構成的意味を計算

意味表現



ビーグルが風邪薬を食べたら病気になる

含意関係認識

- 二つの文の間に含意関係が成り立つかどうかを自動認識する自然言語処理技術
- 含意関係: ある文 t_1 が正しいとした時、もう一方の文 t_2 が正しいと言える
 - t_1 : 犬に風邪薬を飲ませると貧血状態に陥る。
→ t_2 : ビーグルが風邪薬を食べたら病気になる。
 - t_1 : 京都駅は京都の新たな観光名所としてにぎわっている。
→ t_2 : 京都駅は京都の観光名所である。
- 意味の同値性が分かる(＝意味理解)
→ ことばが表す情報が正確に分かる

ロボットは東大に入れるか？

- 人工知能で東大入試を突破することを目指すプロジェクト
- 大学入試はいろいろなタイプの問題があり、高度な言語理解・人工知能の面白い題材

問題文

第 4 問

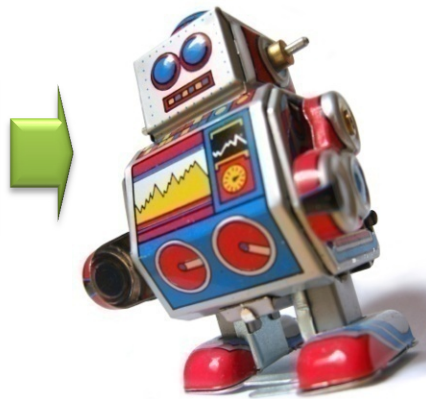
O を原点とする座標平面上の曲線

$$C: y = \frac{1}{2}x + \sqrt{\frac{1}{4}x^2 + 2}$$

と、その上の相異なる 2 点 $P_1(x_1, y_1)$, $P_2(x_2, y_2)$ を考える。

(1) P_i ($i = 1, 2$) を通る x 軸に平行な直線と、直線 $y = x$ との交点を、それぞれ H_i ($i = 1, 2$) とする。このとき $\triangle OP_1H_1$ と $\triangle OP_2H_2$ の面積は等しいことを示せ。

(2) $x_1 < x_2$ とする。このとき C の $x_1 \leq x \leq x_2$ の範囲にある部分と、線分 P_1O , P_2O とで囲まれる図形の面積を、 y_1, y_2 を用いて表せ。



解答

(1) $P_i(x_i, y_i)$ を通る x 軸に平行な直線と、 $y = x$ との交点は $H_i(y_i, y_i)$ であるから、

$$\triangle OP_iH_i = \frac{1}{2}|(y_i - x_i)y_i|$$

ここで、

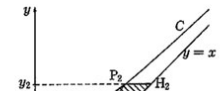
$$\begin{aligned}(y_i - x_i)y_i &= \left(\frac{1}{2}x_i + \sqrt{\frac{1}{4}x_i^2 + 2} - x_i\right)\left(\frac{1}{2}x_i + \sqrt{\frac{1}{4}x_i^2 + 2}\right) \\ &= \frac{1}{4}x_i^2 + 2 - \frac{1}{4}x_i^2 \\ &= 2 \cdots \textcircled{1}\end{aligned}$$

よって、 $\triangle OP_iH_i = \frac{1}{2} \cdot 2 = 1$ ($i = 1, 2$) であるから、 $\triangle OP_1H_1$ と $\triangle OP_2H_2$ の面積は等しい。

$$(2) \frac{1}{2}x + \sqrt{\frac{1}{4}x^2 + 2} > \frac{1}{2}x + \frac{1}{2}|x| \geq 0$$

よって、 C は $y > 0$ を満たす領域にあるから、

① より、曲線 C の方程式は



知識を問う問題を解く

- 知識を問う問題＝教科書や参考書を見れば答えられる問題
- つまり、教科書や参考書の内容をどれだけ記憶しているかが問われる
 - －いわゆる暗記問題
- いろいろなタイプの問題の中で、コンピュータが一番得意そうな問題
 - －だれかがどこかに答えを書いている
 - －つまり、検索の延長線上

知識を問う問題

2009年度センター試験
世界史A

問 3 下線部②の国について述べた文として最も適当なものを、次の①～④のうちから一つ選べ。 3

オスマン帝国について述べた文として正しいものを一つ選べ。

- ① スレイマン1世の時代が最盛期であった。
- ② バルカン半島に誕生した後、小アジアへ進出した。

教科書を見れば...

2009年度センター試験 世界史A

問 3 下線部②の国について述べた文として最も適当なものを、次の①～④のうちから一つ選べ。 3

オスマン帝国について述べた文として正しいものを一つ選べ。

- ① スレイマン1世の時代が最盛期であった。
- ② バルカン半島に誕生した後、小アジアへ進出した。

スレイマン1世

... 46年の在位の中で13回もの対外遠征を行い、数多くの軍事的成功を収めてオスマン帝国を最盛期に導いた。

オスマン帝国

... アナトリアの片隅に生まれたオスマン朝は、やがてバルカン半島など地中海世界の過半を覆い尽くす世界帝国に発展した。

コンピュータは記憶が得意

- コンピュータにとって教科書や参考書の記憶は簡単
 - SDカードやUSBメモリ1個あれば、全ての教科書、参考書の内容を丸ごと記憶できる



知識を問う問題は
コンピュータなら簡単に解けるはず

コンピュータは~~記憶~~が得意 丸暗記

- コンピュータが得意な問題：
 - 世界史Bの教科書28ページの3行目には何と書いてあるか？
 - 一字一句正確に答えられる



人間にとっては無意味な問題

- 人間が記憶しているのは言葉そのものではなく、言葉が表す情報である

||

意味, 知識

知識を問う問題→含意関係認識

- 記憶していることと問われていることが**意味的に合致しているかどうか**を認識する
→ 含意関係認識

オスマン帝国について述べた文として正しいものを一つ選べ。

- スレイマン1世の時代が最盛期であった。
- バルカン半島に誕生した後、小アジアへ進出した。

小アジア
← アナトリアの片隅

スレイマン1世の時代が最盛期
← スレイマン1世が最盛期に導いた

スレイマン1世

... 46年の在位の中で13回もの対外遠征を行い、数多くの軍事的成功を収めてオスマン帝国を最盛期に導いた。

オスマン帝国

... アナトリアの片隅に生まれたオスマン朝は、やがてバルカン半島など地中海世界の過半を覆い尽くす世界帝国に発展した。

NTCIR

- 国立情報学研究所が主催している国際ワークショップ
 - 情報アクセス技術の評価のための共有データを提供
 - 国内外の研究グループが同じデータを用いて評価を行い、知見を共有する
 - 毎回、複数のタスクを設定
 - 特許翻訳
 - 地理時間情報検索
 - RITE: 含意関係認識
- ← 今回、これに参画

センター試験にチャレンジ

- センター試験の問題の一部を含意関係認識で解く
- 対象科目：社会系6科目
 - 世界史A・B, 日本史A・B, 政治経済, 現代社会
- 6チームが参加
 - IBM東京基礎研究所
 - CMU(カーネギーメロン大学)
 - 京都大学
 - 東北大学
 - 北陸先端大学院大学
 - JUCS (Jadavpur University)

データの作り方

センター試験

問 3 下線部②の国について述べた文として最も適当なものを、次の①～④のうちから一つ選べ。 3

- ① スレイマン1世の時代が最盛期であった。
- ❌ 国教はシーア派のイスラーム教であった。
- ❌ バルカン半島に誕生した後、小アジアへ進出した。
- ❌ ベルリン会議により、ボスニア＝ヘルツェゴヴィナの統治権を得た。



Wikipedia

スレイマン1世

この記事は検証可能な参考文献や出典が全く示されていないか、不十分です。
出典を追加して記事の信頼性向上にご協力ください。このタリは2012年2月に貼付けられました。

スルタン・スレイマン1世 (Kanuni Sultan Süleyman, オスマン語: سليمان, Sulaymān, トルコ語: Süleyman, 1494年11月6日 - 1566年9月5日) は、オスマン帝国の第10代皇帝 (在位: 1520年 - 1566年)。

46年の長期にわたる在位の中で13回もの対外遠征を行い、数多くの軍事的成功を収めてオスマン帝国を最盛期に導いた。 英雄では「壮麗帝 (the Magnificent)」のあだ名で呼ばれ、日本ではしばしば「スレイマン大帝」と称される。トルコでは法典を編纂し帝国の制度を整備したことから「立法帝 (カーヌーニ al-Qānūnī / Kanunī)」のあだ名で知られている。

名前のスレイマン (Süleyman) とは、ユダヤ教やキリスト教と共にイ

出典: Wikipedia「スレイマン1世」<http://ja.wikipedia.org/wiki/スレイマン1世>

含意関係: あり

t_1 : スレイマン1世は、46年の在位の中で13回もの対外遠征を行い、数多くの軍事的成功を収めてオスマン帝国を最盛期に導いた。

t_2 : オスマン帝国は、スレイマン1世の時代が最盛期であった。

結果

- センター試験の社会系科目の約半数が含意関係認識に帰着できる
- 現在最先端の含意関係認識技術により、知識を問う問題の5割～6割程度に正答できる
 - えんぴつを転がす(正答率25%)よりだいぶ良い
 - 東大入学レベルにはまだ遠い
- 含意関係認識の精度が向上すれば、試験の正答率も向上すると期待される

コンピュータにとっての 「知識を問う問題」

このチャレンジ
はここ

知識と一致するか
認識

人間 😊
コンピュータ ☹️

試験問題

問 3 下線部②の国について述べた文として最も適当なものを、次の①～④のうちから一つ選べ。 3

- ① スレイマン 1 世の時代が最盛期であった。
- ② 国教はシーア派のイスラーム教であった。
- ③ バルカン半島に誕生した後、小アジアへ進出した。
- ④ ベルリン会議により、ボスニア = ヘルツェゴヴィナの統治権を得た。



知識を問う問題は
元々ここを測る

記憶

人間 ☹️
コンピュータ 😊

教科書・参考書

ウィキペディア フリー百科事典

スレイマン1世

この記事は検証可能な参考文献や出典が全く示されていないか、不十分です。
出典を追加して記事の信頼性向上にご協力ください。このタスクは2012年2月に貼り付けられました。

スルタン・スレイマン1世 (Kanuni Sultan Süleyman, オスマン語: سُليمان، Sulaymān, トルコ語: Süleyman, 1494年11月6日 - 1566年9月5日) は、オスマン帝国の第10代皇帝 (在位: 1520年 - 1566年)。

46年の長期にわたる在位の中で13回もの対外遠征を行い、数多くの軍事的成功を収めてオスマン帝国を最盛期に導いた。英語では「**壮麗帝** (the Magnificent)」のあだ名で呼ばれ、日本ではしばしば「**スレイマン大帝**」と称される。トルコでは法典を編纂し帝国の制度を整備したことから「**立法帝** (カーヌーニ al-Qānūnī / Kanuni)」のあだ名で知られている。

名前のスレイマン (Süleyman) とは、ユダヤ教やキリスト教と共にイ

スレイマン1世



RITE2 を企画中

- 詳しい情報は
 - <http://www.cl.ecei.tohoku.ac.jp/rite2>
 - Twitter: @NTCIR10_RITE
- 人間が試験を解くプロセスにより近い設定
 - 根拠となる文章を明示的に与えず、Wikipediaや教科書から自分で探す
- スケジュール
 - 参加申込締切: 6月30日
 - データ公開: 7月1日
 - 最終評価: 11月中旬
- 興味のある方はチャレンジしてください

含意関係認識はなぜ難しいのか？

- たくさんの研究にもかかわらず、含意関係認識の精度はいまだに低い
 - 50%～70%程度
- 構成的意味・語彙的意味を計算すれば、含意関係認識ができるはずでは？

構成的意味と語彙的意味の相互作用

- 語彙的意味：単語の置き換え可能性を表す
- **いつでも置き換えられるわけではない**
 - 構成的意味の構造により、置き換え可能性が変わる

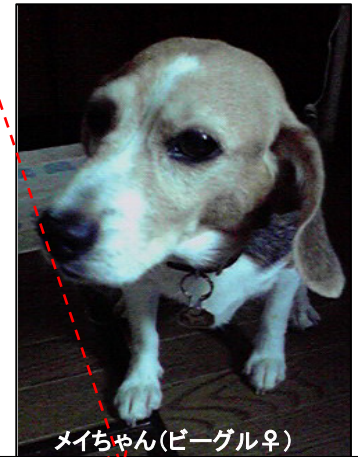
犬に風邪薬を飲ませると貧血状態に陥ります。

下位

類義

上位

含意



メイちゃん(ビーグル♀)

ビーグルが風邪薬を食べたら病気になる！

構成的意味と語彙的意味の 相互作用

- 犬に風邪薬を飲ませると貧血状態に陥ります。
- ✗ → 動物が風邪薬を食べたら貧血になる
- 犬に風邪薬を飲ませると貧血状態に陥ることがあります。
- → 動物が風邪薬を食べたら貧血になるかもしれない
- 犬に風邪薬を飲ませるのをやめた。
- ✗ → ビーグルに風邪薬を飲ませるのをやめた。

しくみはまだ完全には分かっていない

シソーラス・オントロジーの限界

- なぜ「スレイマン1世が最盛期に導いた → スレイマン1世の時代が最盛期」と言えるのか？
- 現在の技術では、同義フレーズパターンとして扱うしかない
- フレーズは無限にある

オスマン帝国について述べた文として正しいものを一つ選べ。

- ① スレイマン1世の時代が最盛期であった。
- ② バルカン半島に誕生した後、小アジアへ進出した。

スレイマン1世の時代が最盛期
← スレイマン1世が最盛期に導いた

スレイマン1世

... 46年の在位の中で13回もの対外遠征を行い、数多くの軍事的成功を収めてオスマン帝国を最盛期に導いた。

意味に関わるその他の問題

- 時間、アспект、様相
 - お風呂が沸きました
 - お風呂が沸いています
 - お風呂を沸かしました
 - お風呂を沸かしています

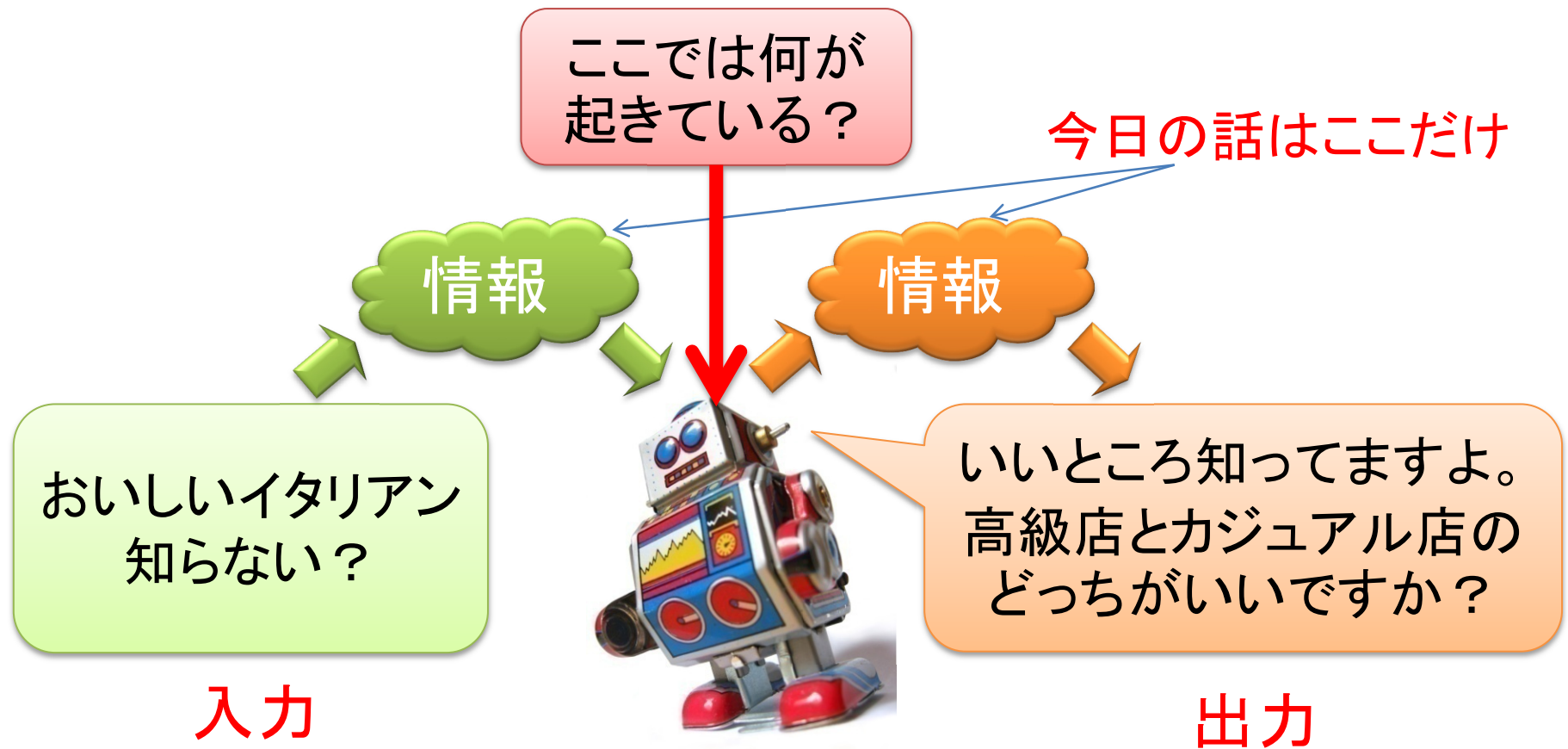
→ お風呂に入れる

→ まだお風呂に入れない
- 参照関係
 - (誰が?) 46 年の在位の中で13 回もの対外遠征を行い、数多くの軍事的成功を収めて...
- メタファー、メトニミー
 - 太宰治を読んだ。
 - 官邸は首脳会談の日程を発表した。

これらの問題を解決してはじめて、
人間と同等の意味理解が可能になる

「意味理解」から「言葉の理解」へ

- ここまでの話は、「意味理解」の話



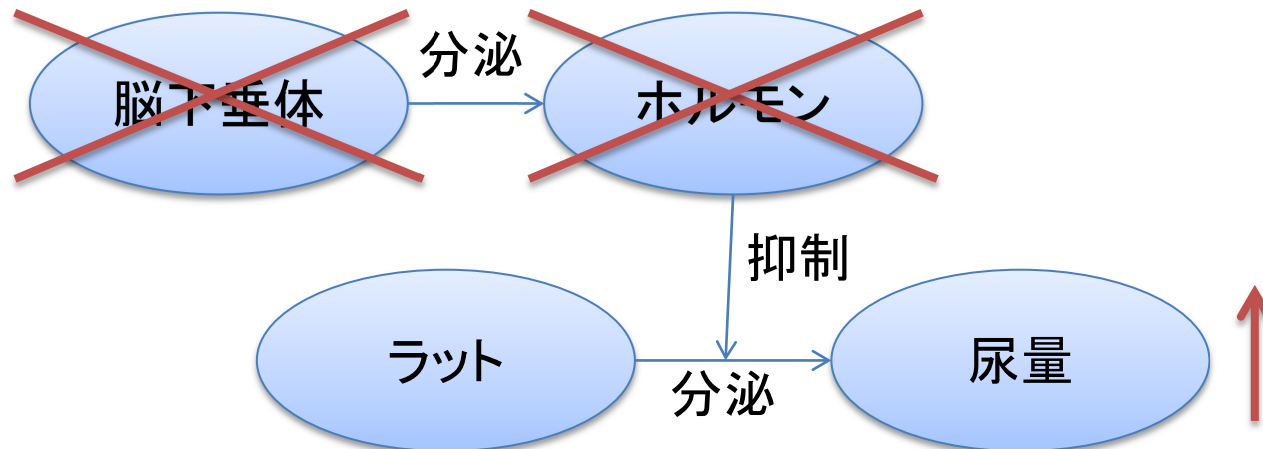
定性的推論

2009年度センター試験 生物 I

脳下垂体からは様々なホルモンが分泌される。それらの作用を調べるために、ラットを麻酔し、苦痛のない状態で脳下垂体の摘出手術を行い、その後の様子を観察した。脳下垂体を摘出した後、ラットに起こる変化として最も適当なものを選べ。

1. 尿量が増加する。
2. 代謝が盛んになる。

• 知識を検索するだけでは解けない



自然な状況の認識

2009年度センター試験 英語

問 1 Zack

もう10時だ。ボブがトイレから帰って来たら出た方がいいね。割り勘にする？

Koji

いや。僕は君たちよりたくさん食べて飲んだよ。僕は多く払うべきだと思う。

Zack:

18

Koji

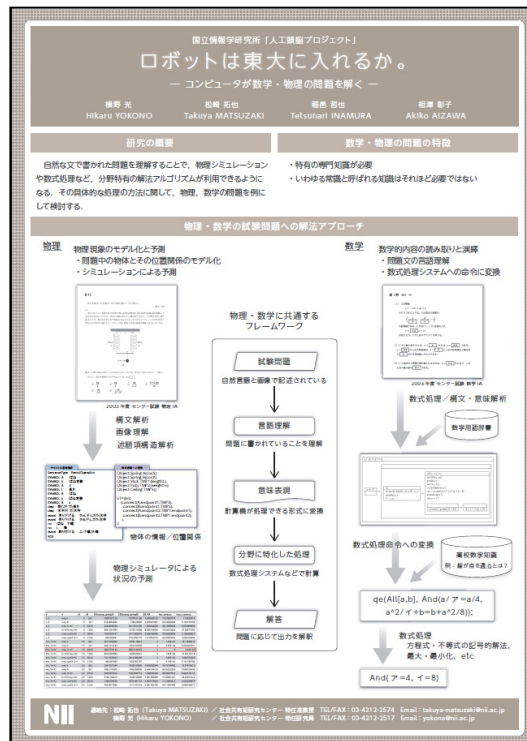
それは公平だね。

- ① 落ち着いて。そんなに興奮すること無いよ。
- ② まけてくれるよう頼んでみる？
- ③ 雑誌のクーポンを持ってくればよかった。
- ④ 別々に払うようにする？

- どうして「自然な状況」が分かるのか？
- 経験、常識に基づく判断

その他もろもろ...

- 試験問題を分析すると、言語理解や人工知能のさまざまな問題が見えてくる
- ポスター展示へどうぞ



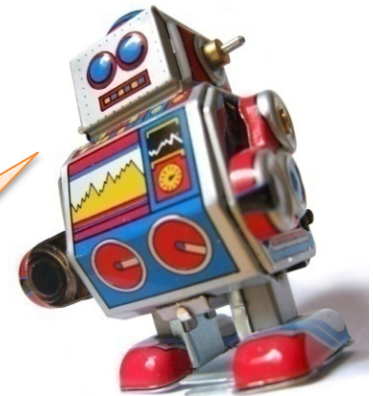
おわりに

- 自然言語の意味の同値性・差異を計算するためにはいろいろな理論・技術が必要
 - 構文解析、シソーラス・オントロジー、あいまい性解消、...
- 現在最先端の含意関係認識で、センター試験をある程度解くことができる
- 人間が理解するように「ことばを理解する」にはまだ遠い
- 人間にとって難しいこととコンピュータにとって難しいことは異なる



おいしいイタリアン
知らない？

いいところ知ってますよ。
高級店とカジュアル店の
どっちがいいですか？



ご清聴ありがとうございました