

お茶大での博士課程の 生活について

OCHADAI OPEN CAMPUS 2025

博士後期課程 1年 富田朝

2025 / 07 / 19

1. はじめに

今日お話しすること

1

1. 自己紹介・情報科で学べる内容（論理学）
2. 研究内容について
3. なぜ博士課程に進学したのか？

1. はじめに

自己紹介 – 基本プロフィール

2



名前 **富田 朝** (とみた あさ / Asa Tomita)

所属 人間文化創成科学研究科 理学専攻 情報科学領域
博士後期課程1年 戸次研究室

研究分野 **数理論理学**・計算言語学・自然言語処理

出身 埼玉県さいたま市

趣味 バスケット、テトリス、猫、お昼寝、アニメ

連絡先 ホームページ : <https://morning85.github.io/>
メールアドレス : tomita.asa@is.ocha.ac.jp

1. はじめに

自己紹介 - 経歴

3



2016
～ 2019

高校

お茶の水女子大学附属高等学校

2019
～ 2023

学部

お茶の水女子大学
理学部 情報科学科

2023
～ 2025

修士

お茶の水女子大学大学院
人間文化創成科学研究科 理学専攻 情報科学コース

2025
～ 現在

博士

お茶の水女子大学大学院
人間文化創成科学研究科 理学専攻 情報科学領域

Q. 前提1,2から帰結はどのように証明できるのか

前提1. 全ての人間は死ぬ

前提2. ソクラテスは人間である

仮説. ソクラテスは死ぬ

$P(x)$: x は人間である

$Q(x)$: x は死ぬ

と定義するとき、前提1, 2と帰結は以下のように表せる

前提1. $\forall x. (P(x) \rightarrow Q(x))$

前提2. $P(\text{ソクラテス})$

仮説. $Q(\text{ソクラテス})$

情報科学科で学べること - 一階述語論理

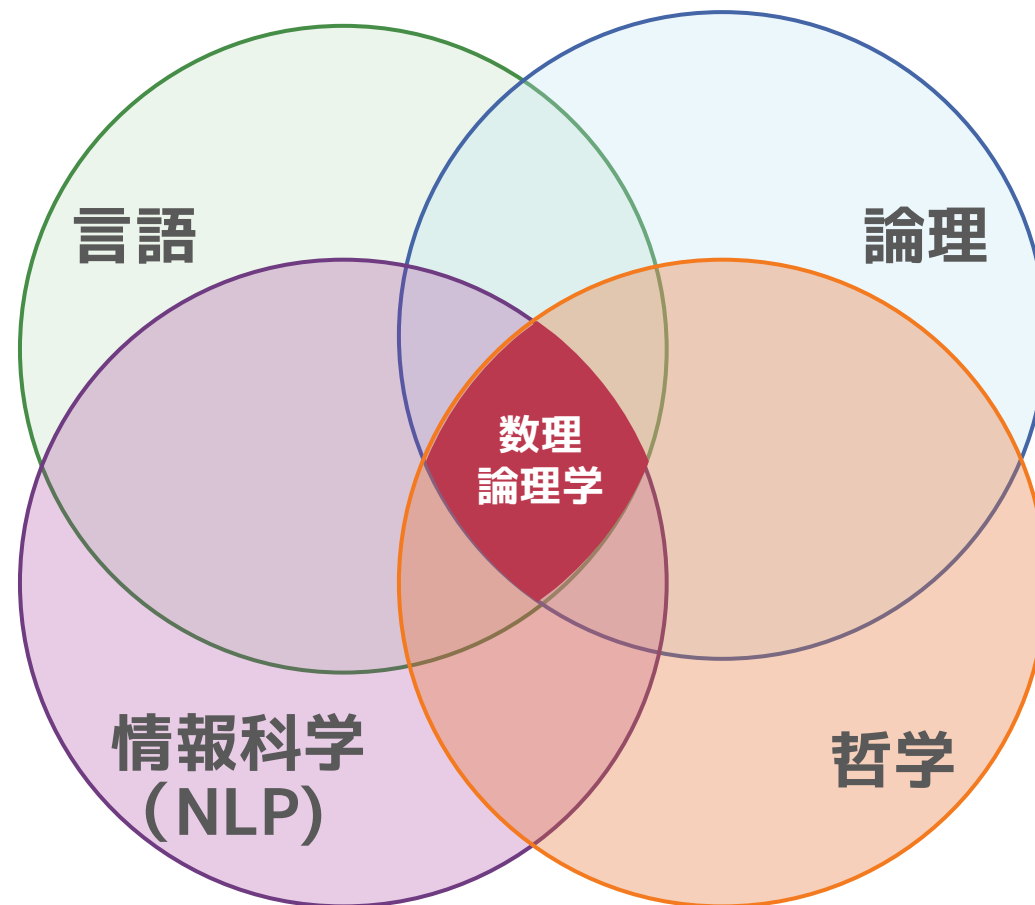
$$\forall x. (P(x) \rightarrow Q(x)), P(\text{ソクラテス}) \models Q(\text{ソクラテス})$$

$$\begin{array}{c} (\forall E) \frac{\forall x. (P(x) \rightarrow Q(x))}{P(x) \rightarrow Q(x) [\text{ソクラテス} / x]} \\ \equiv P(\text{ソクラテス}) \rightarrow Q(\text{ソクラテス}) \quad P(\text{ソクラテス}) \\ (\rightarrow E) \frac{}{Q(\text{ソクラテス})} \end{array}$$

研究分野

- 数理論理学
- 理論言語学
- 計算言語学

→ 言語学、論理学、情報科学、哲学の融合分野



計算言語学とは

研究課題

- 言語の意味とは何か
- 言葉が通じる仕組みは、どのように「計算」されるのか

(1) 富士山の頂上には年中スターバックスが営業していて、登山客はどんな天候でも温かいコーヒーが飲める。

- 私たちは、初めて見た文でも「意味」が理解できる
しかし、「文の意味」が理解できることと文の意味が「真」であることは全く別の問題
- 文の意味を記述することができれば、文間の関係性について考えることができる

言語の意味への2つのアプローチ

自然言語処理

使用説

- 語の意味は言語におけるその使われ方のこと

分布意味論

統計とベクトル表現で意味を表現する

分布仮説：語の意味はその後の周辺に現れる語（文脈）によって定まる

計算言語学

真理条件説

- 文の意味とは、真理条件（どのような状況では真であり、どのような状況では偽となるか）のこと

形式意味論

論理と記号表現で意味を表現する

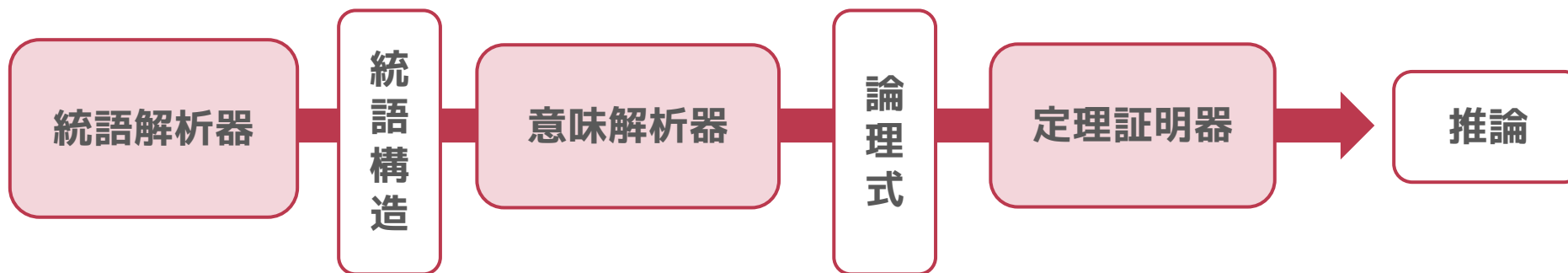
計算言語学の推論

言語の意味が形式的に記述できると、**推論**（ある命題から別の命題を導くこと）ができるようになる

例：含意関係認識

前提文：太郎が走った

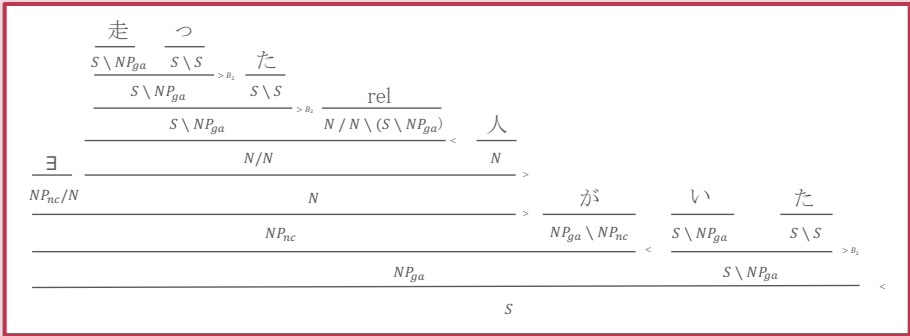
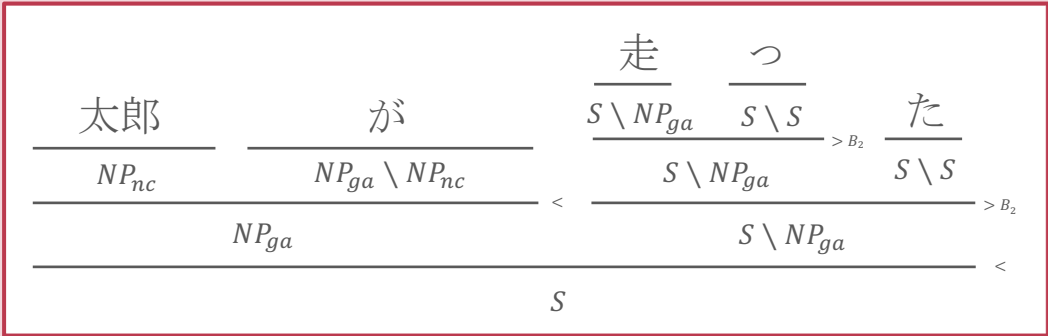
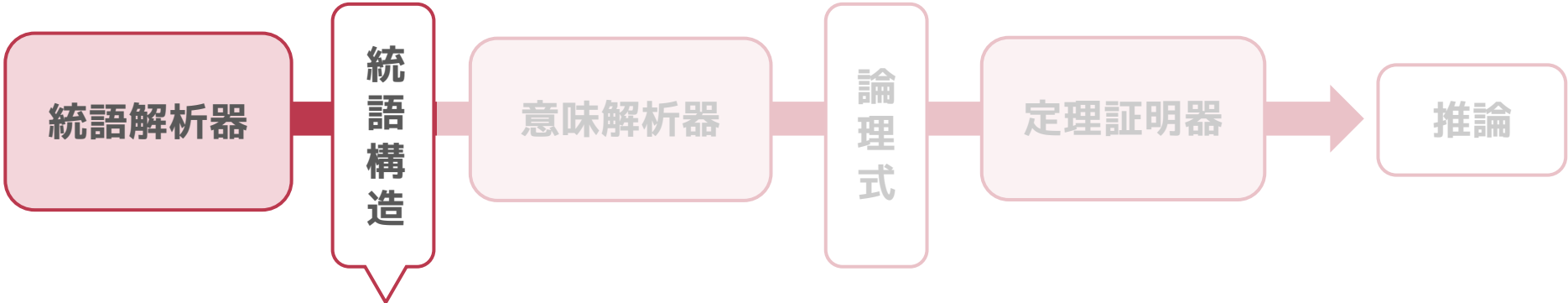
仮説文：走った人がいた



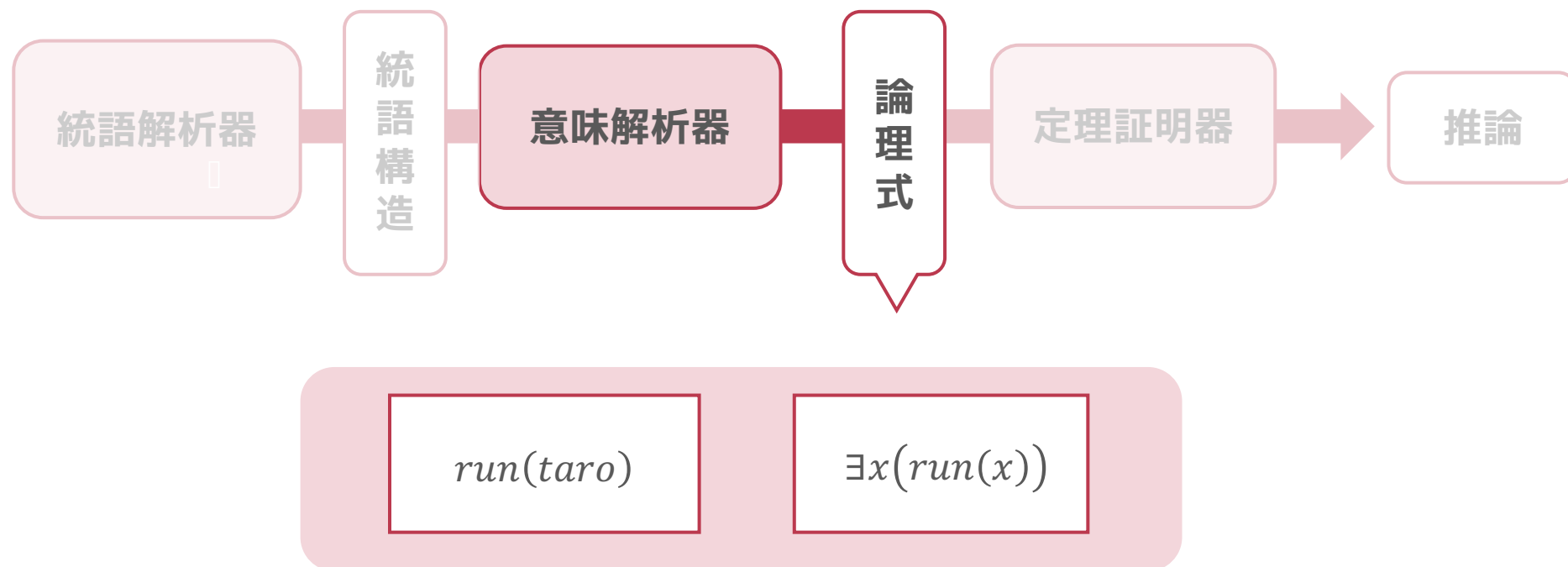
2. 研究内容

統語解析

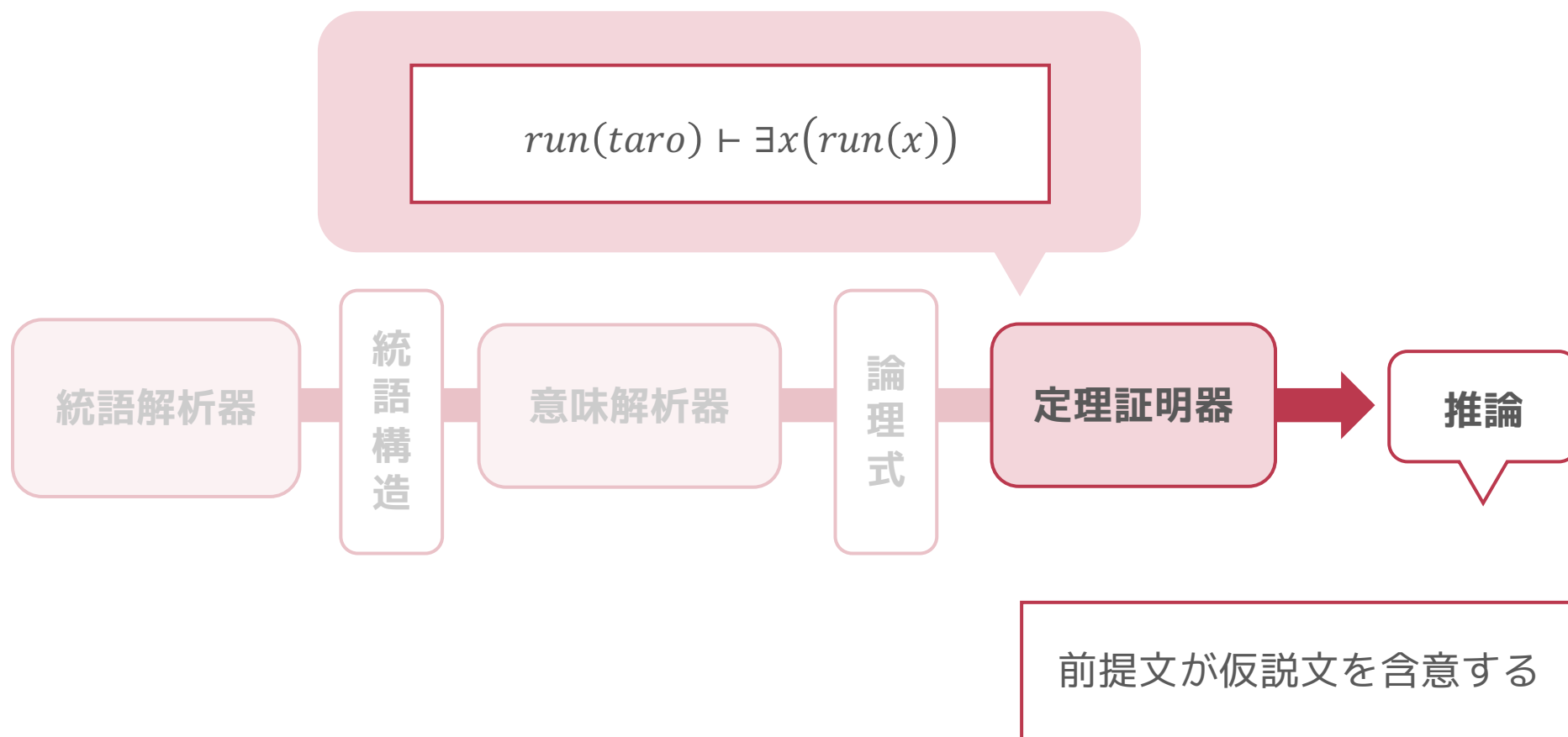
前提文：太郎が走った
仮説文：走った人がいた



前提文：太郎が走った
仮説文：走った人がいた



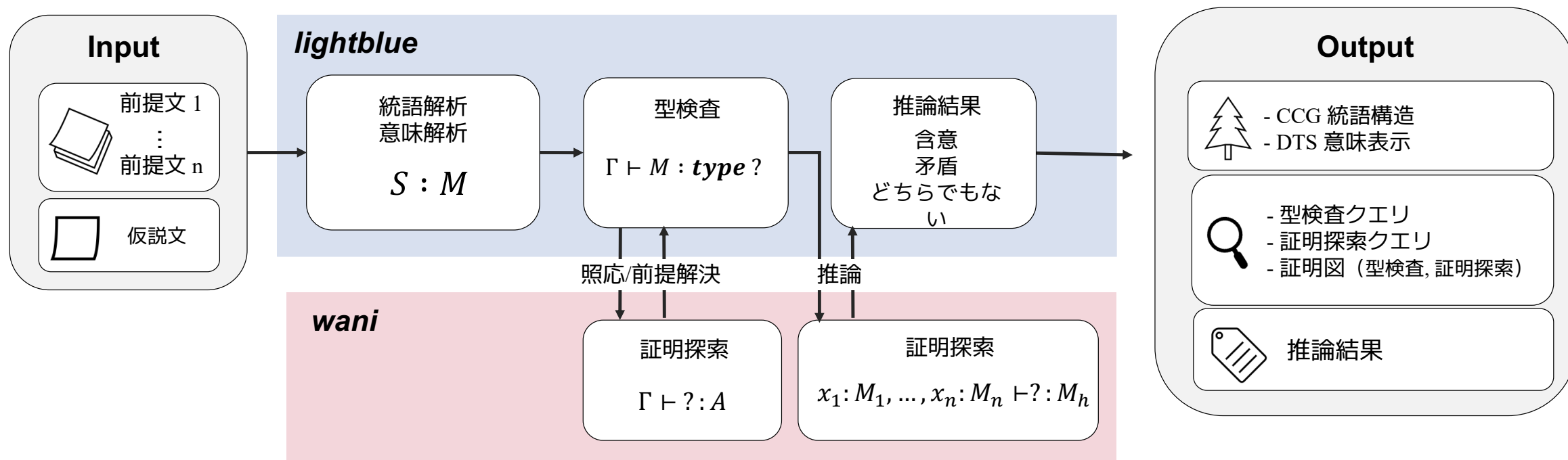
定理証明・推論



2. 研究内容

研究テーマ1: 推論システムの改良

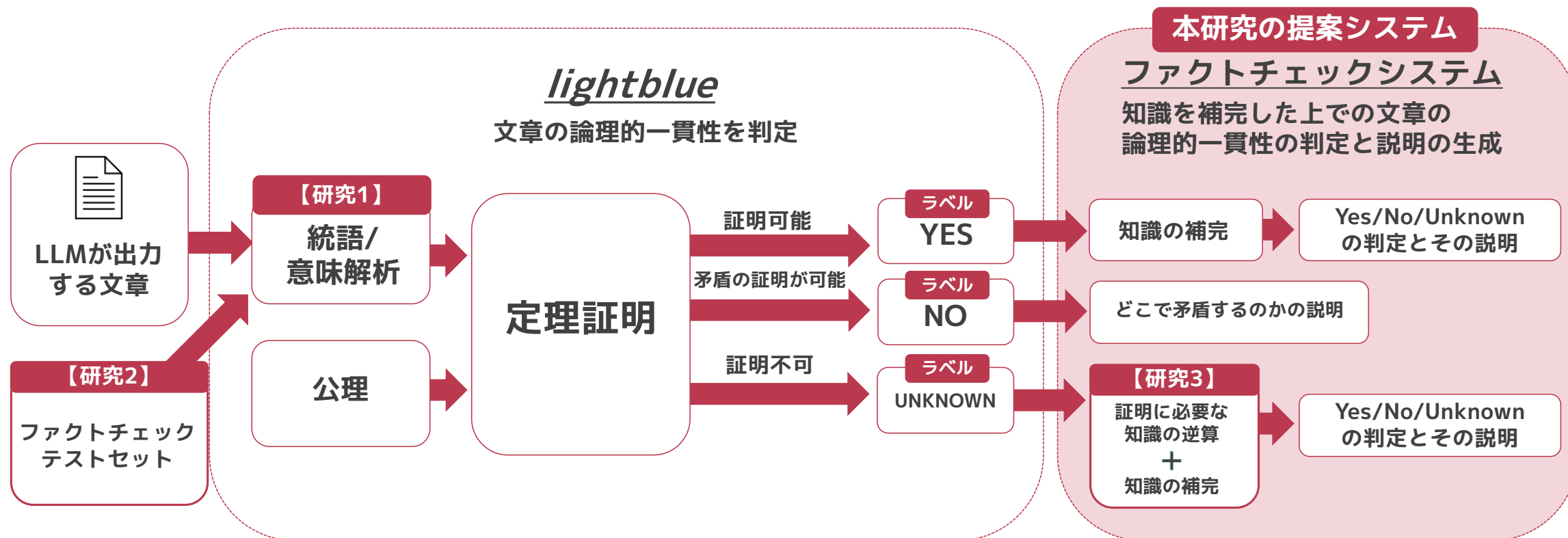
13



2. 研究内容

研究テーマ2: ファクトチェックシステムの構築

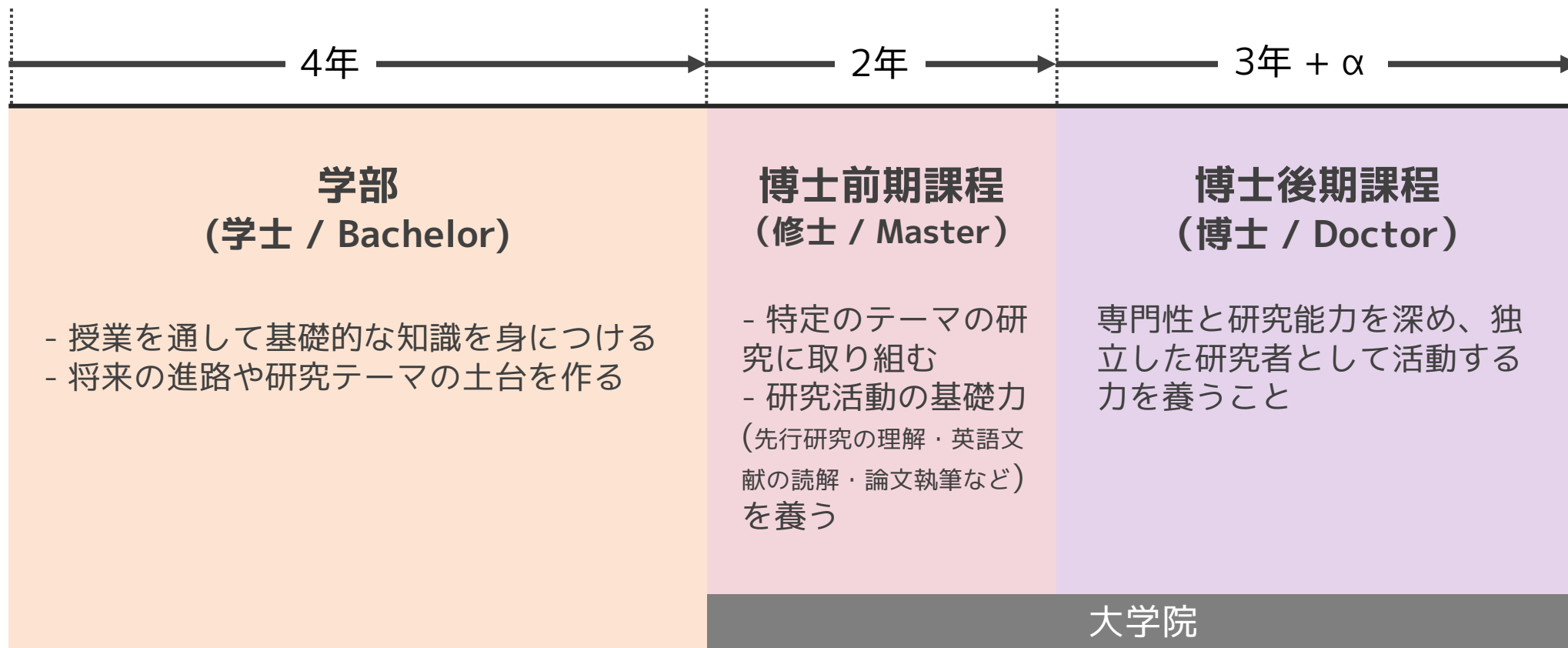
14



3. なぜD進したのか

博士後期課程とはどんな場所？

15



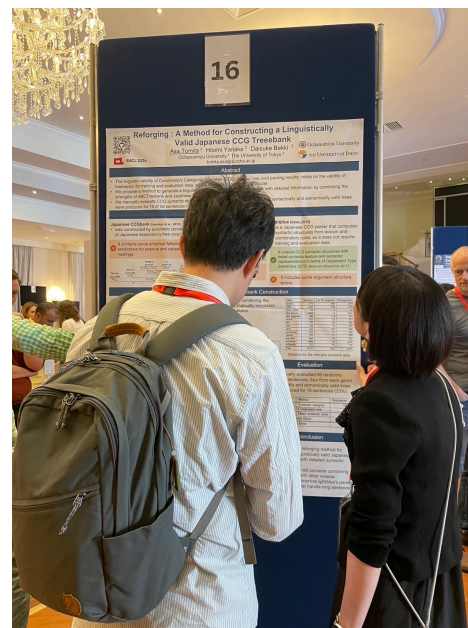
3. なぜD進したのか

進学のきっかけ1: 国際学会

16

修士1年で国際学会に参加し、研究のモチベーションがあがった

- 研究を続ければ、グローバルに活躍できることを実感した
- 同世代の研究での活躍に刺激を受けた



Reforging: A Method for Constructing a Linguistically Valid Japanese CCG Treebank

Asa Tomita¹ Hitomi Yanaka² Daisuke Bekki¹

¹Ochanomizu University, Japan

²The University of Tokyo, Japan

{tonita.asa,bekki@is.ocha.ac.jp}

hyanaka@is.s.u-tokyo.ac.jp

Abstract

The linguistic validity of Combinatory Categorical Grammar (CCG) parsing results relies heavily on treebanks for training and evaluation, so the treebank construction is crucial. Yet the current Japanese CCG treebank is known to have inaccuracies in its analyses of Japanese syntactic structures, including passive and causative constructions. While ABCTreebank, a treebank for ABC grammar, has been made to improve the analysis, particularly of argument structures, it lacks the detailed syntactic features required for Japanese CCG. In contrast, the Japanese CCG parser *lightblue*, efficiently provides detailed syntactic features, but it does not accurately capture argument structures. We propose a method to generate a linguistically valid Japanese CCG treebank with detailed information by combining the strengths of ABCTreebank and *lightblue*. We develop an algorithm that filters *lightblue*'s lexical items using ABCTreebank, effectively converting *lightblue* output into a linguistically valid CCG treebank. To evaluate our treebank, we manually evaluate CCG syntactic structures and semantic representations and analyze conversion rates.

1 Introduction

There have been significant advances in natural language processing research through the construction of syntactic tree corpora, known as treebanks. Treebanks are datasets where syntactic structures are annotated over large bodies of text. Various treebanks (Marcus et al., 1993; Foris, 2003; Briscoe and Carroll, 2006; Hockenmaier, 2006; Hockenmaier and Steedman, 2007; Vadas and Curran, 2007; Ito et al., 2010; Boswell and Brew, 2010) have been served as standard datasets for training and evaluating statistical syntactic parsers. The Penn Treebank (Marcus et al., 1993), one of the first context-free grammar (CFG) treebanks, contains a one-million-word corpus of Wall Street Journal text. CCGbank (Hockenmaier and Steed-

man, 2007) was constructed by converting the Penn Treebank to Combinatory Categorical Grammar (CCG; Steedman, 1996, 2000), which contributing to the advancement of CCG parsers.

There are various methods for constructing treebanks. One approach is to combine automatic part-of-speech (POS) taggers and syntactic parsers with manual corrections, as in the Penn Treebank. The approach to providing CCGbank involves automatic conversion from existing treebanks. However, treebanks have different formats, provided information, and informational validity. Japanese CCGbank (Uematsu et al., 2013), constructed by automatic conversion of Japanese dependency tree corpora, but provides limited validity of syntactic structures for passive or causative sentences (Bekki and Yanaka, 2023). The Japanese ABC grammar treebank ABCTreebank (Kubota et al., 2020) has manually annotated argument structures, but does not provide POS information (conjugation series, conjugation forms, among others) and other detailed information.

We thus aim to construct a Japanese CCG treebank with both linguistically valid syntactic structures and detailed syntactic features. To this end, we propose a method to construct a new Japanese CCG treebank using the Japanese CCG parser *lightblue* (Bekki and Kawanoe, 2016)¹, which can output detailed syntactic features. However, *lightblue* contains errors related to argument structures, causing inaccurate outputs. To address this drawback, we extracted predicate-argument structures from ABCTreebank and incorporated this information into *lightblue*. This approach involves decomposing and reconstructing the treebank, which we call "reforging". We discuss the reforging process in more detail in Section 3. Section 4 assess the validity of our proposed method. Section 5 introduces the error analysis of the output trees. Section 6

196

進学のきっかけ2: 将来のキャリア

就活を通して、将来研究職に就きたいと思うようになった

- 研究職は博士号を持っていた方が有利
- 社会人をやりながら博士課程で研究も進める社会人Dでは、博士号取得までに時間がかかりそうに感じた

理系博士（特に情報系）は重宝されるため、博士号を取得することで、自分のキャリアの可能性を広げることができる

進学のかっかけ3: 競争的資金の確保

お茶大生が申請できる（富田が申請した）研究奨励金のプログラム

1. 日本学術振興会特別研究員（DC1/DC2） 不採用
 - 奨励金：20万/月
 - 研究費：最大150万/年
2. 次世代研究者挑戦的研究プログラム（JST SPRING） 採用（辞退）
 - 奨励金：20万/月
 - 研究費：20万/年
3. 次世代AI人材育成プログラム（JST BOOST） 採用
 - 奨励金：25万/月
 - 研究費：90万/年
4. 大学院博士後期課程研究奨励賞 不採用
 - 10万/年

進学してよかったこと

- とにかく研究に集中できる
 - 修士では、就活があり、研究に100%で打ち込める時間が短かった
 - 修士よりも自分で研究の舵を切れるので、自由度が上がった
- 海外での研究発表のチャンスが増えた
 - 8月にドイツ（Bochum）・スロベニア（Ljubljana）、9月にドイツ（Düsseldorf）で研究発表をする予定
- （学割が使える）
- （時間の融通が利きやすい）

3. なぜD進したのか

お茶大の博士課程進学について

20

お茶大の情報科学領域は、博士課程支援が充実してきている

- 大学・高専機能強化支援事業に採択
- 研究支援プログラム (SPRING, BOOST)に採択
- 定員が7名→12名に増枠

情報科学領域の**博士後期課程進学**の**ススめ**が公開されています。覗いてみてください。

<https://www.is.ocha.ac.jp/to-phd/>

Department of Information Sciences, Ochanomizu University Japanese only

情報科学科	博士前期課程 情報科学 コース	博士 後期課程 情報科学 領域	進路状況	公募採択 受賞記録	入試情報
-------	-----------------------	--------------------------	------	--------------	------

[HOME](#) > 博士後期課程進学のススめ

博士後期課程進学のススめ

本学の理工学系情報科学領域では博士後期課程に進学する人を熱烈に歓迎いたします。ほんの1%でも興味のある方はぜひ本ページを一通り読んでいただければと思います。

博士後期課程に進学する意義

博士後期課程への入学は研究を続けるための第一歩であるだけでなく、皆さんの人生を大きく広げる分岐点でもあります。博士後期課程への進学によって、以下に示すような素晴らしい未来が皆さんを待っています。

- ・ **研究を続ける意義**
まだ解明されていない学問の課題を解くための一員となり、世界の誰も解明していない新しい知見を見出す一員となります。研究という工程が少しでも好きな人にはこれ以上ない宝庫の時間が待っています。そして生涯にわたって有用となる高度な問題発見能力や問題解決能力を身に付ける機会にもなります。
- ・ **世界に羽ばたく**
世界的な学会で発表することで、学問社会の世界の一員になります。世界中の研究者に論文を読まれるようになり、世界中に優秀な仲間を作ることが可能になります。
- ・ **世界に貢献する**
新しい研究成果を打ち出すことで、学問の新しい流れを作る一員になります。また、社会を大きく変える新しい技術や、社会問題解決に貢献する貴重な技術を開発することが可能になります。
- ・ **専門性の高い人生を切り拓く**
大学教授などのアカデミックな職業だけでなく、企業の研究員なども含めて、職業として研究者になるためには博士号はほぼ必須な状況となっています。また企業でも研究職以外の多くの専門的な職種において、博士号を取得した多くの人が活躍しています。さらにいうと、世界の主要国の一流の企業人には、研究職以外の職種の人にも博士号取得者が非常に多くいます。このような世界の一流の企業人と対等に見られるためにも、博士号取得が有効に働くことがあります。
- ・ **修士まででは研究は終わらない**
博士前期課程に進学した皆さんは、学部卒業研究から通すと一定年数の研究を経験したことになります。しかし研究とは多くの勉強と経験を重ねる工程であり、修士までの研究でそれが完了するとは言えないものがあります。ましてや入学するやいなや就職活動に大きな時間を費やしている修士学生にとって、博士前期課程で研究に没頭できる時間はとても十分なものとさえ思いません。修士までの研究で皆さんは本当に「やりきた」と言えますでしょうか。本格的に「研究とは何か」「学問とは何か」を考えてみて、修士までの研究を「博士後期課程で活躍するための助走」と考えてみませんか？

博士後期課程進学後のキャリア

本学情報科学領域の博士後期課程進学後のキャリアは非常に多岐で、むしろ学術機関から企業から引っ張りだこな状態が続いています。博士後期課程に進学したら担って役割が重くなるのでは...といった話を聞くことがあるかもしれませんが、本学情報科学領域では全くその心配はありません。以下に代表的な進路を紹介します。

- ・ **大学・研究所**
研究者として大学や研究所などの学術機関に就職するには博士号の取得は必須です。進学を送る必要はありません。