

自然言語における知識獲得及び常識推論 への接近法

-移民政策と日本語教育、法的な視点から-

2021年11月27日(土) 14:10-14:35

於 人工知能学会 合同研究会2021

第124回 知識ベースシステム研究会 (SIG-KBS)

テーマ:知識表現・知識獲得とその応用

○太田 博三(放送大学, (株)VSN)

目次

1. 研究の背景・目的・新規性
2. 国内外の移住者と第二外国語教育
3. 記号推論の発展と適用の検討
4. 実務における常識の定義と課題の考察
5. 言語的側面からの示唆と検討
6. まとめ

1. 研究の背景・目的・新規性

[先行事例]

1. ドイツやオランダ(第二外国語教育)

- ・対象: 海外からの移民の受入れ
- ・実施者: 国策(移民法)
- ・教育時間: 600時間 + 45時間



[後発事例]

2. 日本(第二外国語教育)

- ・対象: 海外からの労働者
- ・実施者: 自治体 + ボランティア
+ 文化庁国語審議会(2011-)
- ・教育時間: 80時間 + α



3. ねらい・目標

- ・実務としての第二外国語教育
+ 社会・文化の知識獲得



1. 研究の背景・目的・新規性

[先行事例] ⇒ 幼児の言語習得法/DARPA

1. ドイツやオランダ(第二外国語教育)

- ・対象: 海外からの移民の受入れ
- ・教育時間: 600時間 + 45時間



[後発事例] ⇒ 成人の第二外国語の言語習得

2. 日本(第二外国語教育)

- ・対象: 海外からの労働者
- ・実施者: 自治体 + ボランティア
+ 文化庁国語審議会(2007-)
- ・教育時間: 80時間 + α



3. ねらい・目標

- ・実務としての第二外国語教育
+ 社会・文化の知識獲得

ここを本研究
で考察する

⇒ 成否の要因: 時間増? + 場面別? + α ?



1. 研究の背景・目的・新規性

・実務としての第二外国語教育

+ 社会・文化の知識獲得

⇒ 成否の要因: 時間増? + 場面別? + α ?

1. 第二外国語のカリキュラム:

場面別の分類(大分類・中分類・小分類)

⇒ 日常生活を送る上で, 必要な場面.

→ ヒトの言語習得のプロセスを考察する.

2. 1.のヒトの考察を踏まえて, 機械へ適用を検討(対話システムなどへ).

→ ここでは, 記号論理学の帰納法やアブダクション, Multi-Hop Reasoningなどを活用する.

→ 機械ではどのくらい捕捉できるのか考察する.

1. 研究の背景・目的・新規性

標準的なカリキュラム案で扱う生活上の行為			
○ 健康・安全に暮らす			
・ 健康を保つ			
・ 安全を守る			
・ 住環境を整える			
○ 消費活動を行う			
・ 物品購入・サービスを利用する			
・ お金を管理する			
○ 目的地に移動する			
・ 公共交通機関を利用する			
・ 自力で移動する			
○ 人とかかわる			
・ 他者との関係を円滑にする			
○ 社会の一員となる			
・ 地域・社会のルール・マナーを守る			
・ 地域社会に参加する			
○ 自身を豊かにする			
・ 余暇を楽しむ			
○ 情報を収集・発信する			
・ 通信する			

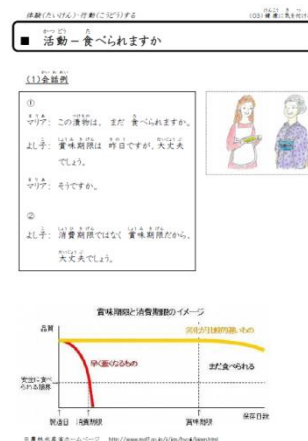


I 標準的なカリキュラム案で扱う生活上の行為の事例			
I 健康・安全に暮らす [7単位]			← 大分類
01 健康を保つ [3.5単位]			← 中分類
(01) 医療機関で治療を受ける [2単位]			← 小分類
・ 隣人に容態を伝えて助言を求める			
・ 初診受付で手続をする			
・ 医者診察を受ける			
・ 病気への対処法・生活上の注意などを質問し答えを理解する			



・標準カリキュラムの場面分類と大分類・中分類・小分類との対応

医療機関で治療を受ける		
	場面	
場所	相手	状況・動機



1. 研究の背景・目的・新規性(に向かうと思われること)

■既存研究では

1. (ヒト対)ヒト: 場面分類が中心であった.

→言語教育での場面分類での知識の限定.

2. (ヒト対)機械: 医療など専門分野, ルールベース

→専門知としての成功と一般的な知識での不成功

■本研究では

(ヒト対)ヒトの場面分類を (ヒト対)機械に適用しようとするもの
場面別での, 語用論や記号推論を適用できる可能性を示唆

2. 国内外の移住者と言語教育

1. ドイツの移民政策と言語教育

カリキュラム構成:

⇒言語教育と社会や文化の教育は分けられている.

⇒シンプルな知識獲得の視点として位置づけられている.

ドイツ語教育で扱うテーマ	
基礎語学:	
①自分自身、②社会的コンタクト、③住居、④買い物と商売、⑤飲食、⑥人間の身体と健康、⑦日常生活、⑧役所や官庁とそのサービス、⑨労働と職業、⑩教育や職業訓練ならびに学習、⑪交通、⑫余暇、⑬自然および環境	
中級語学:	
①メディアおよび現代の情報技術、②社会・国家および国際組織、③他の人々との関係、④文化および世界観	
オリエンテーションコース (30時間)	
①法秩序 (国家の構成、選挙権、州および連邦州、法治国家、社会国家の原則、基本法、住民の義務、官庁や役所の機能)、②歴史 (ドイツ連邦共和国の成立と発展)、③文化 (人間像、時代の理解、規則のオリエンテーション、宗教の多様性)	

第二外国語教育

社会や文化

2.国内外の移住者と言語教育

(参考1) 日本の標準カリキュラムの構成:

⇒6割は読み書きと意思伝達, 残り4割が相互作用を伴う対話と位置づけられている.

大分類	小分類
I 健康・安全に暮らす	01 健康を保つ
	02 安全を守る
II 住居を確保・維持する	03 住居を確保する
	04 住環境を整える
III 消費活動を行う	05 物品購入・サービスを利用する
	06 お金を管理する
IV 目的地に移動する	
VII 人とのかかわる	
	14 他社との関係を円滑にする
VIII 社会の一員となる	15 地域・社会のルール・マナーを守る
	16 地域社会に参加する
IX 自信を豊かにする	20 余暇を楽しむ
	21 通信する
X 情報を収集・発信する	22 マスメディアを利用する

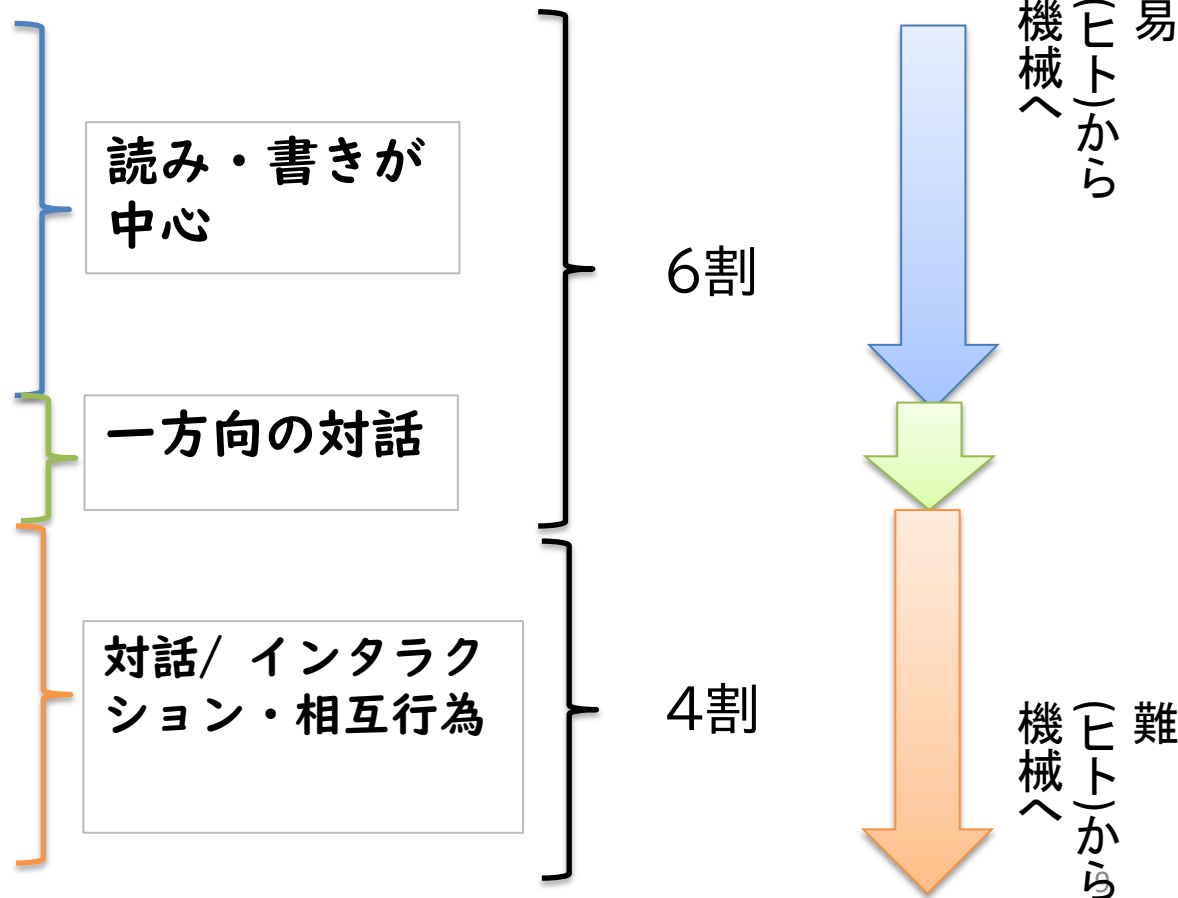


図3-2-2 標準的カリキュラムと学習順序

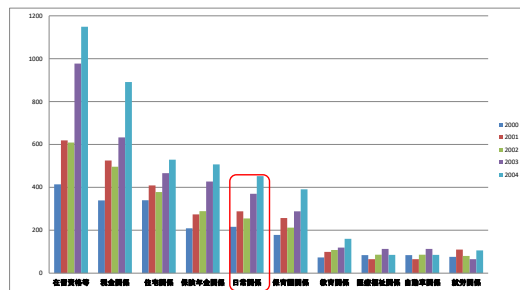
2.国内外の移住者と言語教育

(参考2) 日本の標準カリキュラムの構成:

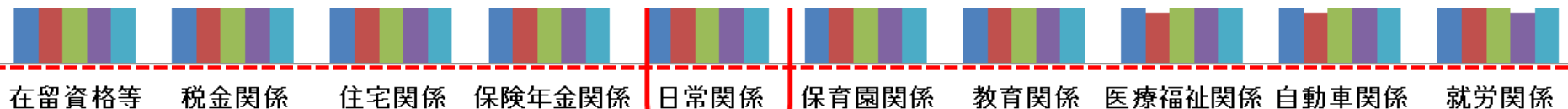
2. 日本のカリキュラムの構成:

→自治体の相談件数の頻度順

→意思伝達や目的を達成するための行為が大半で、生活に必要な場面の大半でもある。



群馬県太田市における外国人相談の内容と相談件数(2000年から2004年)



場面分類の基準に

2.国内外の移住者と言語教育

3.皮肉・アイロニーやポライトネス(≡配慮表現)はコミュニケーション違反!?

協調の原理(ポール・グライス)の4つの公理に反しているときに生じる.
→長期に渡る同一人物間での関係性が前提となり, 省略されて, 対話に現れるもの.

→→ヒトは経験から推測してその含意(本音)を捉える.

表3-2-1 嫁と姑の間の会話断片

A: 姑	「この料理, おいしいわね」
B: 嫁	「そっ、そうですか・・・」

皮肉・アイロニー

表3-2-1 AとBの会話断片 (敬語のない配慮表現)

A	急いている?
B	別に.
A	悪いけど, ペン貸してくれる?
B	いいよ.

ポライトネス(暗示的な敬意や配慮表現)

表3-2-2 相手と状況・動機

A (友人)	B のペンが何本か目の前の机の上に置いてある.
B (友人)	相手B が, 既にペンを筆入れ, さらに鞆にしまっており, 席を立とうとしている.

3.記号推論の発展と適用の検討

3.1.記号論理学による表現と推論

- 機械での知識表現の成功例
 - エキスパートシステム: 医師の診察など
 - 「IF-THEN」形式
 - 1つ1つの論理命題の列挙で構成され, 分岐するものに限る. に限られる.
 - 記号論理学の一階述語論理での表現するが, 数字の単位は表せず, 非常に曖昧なものになる.
- 限界がある.

私はコンピューターを持っている	$\exists x(\text{have}(\text{I}, \text{computer}(x)))$
すべての作家は書物が好き	$\forall x(\text{writer} \rightarrow \exists y(\text{like}(x,y) \wedge \text{documents}(y)))$
図4-1-1一階述語論理式の例	

3.記号推論の発展と適用の検討

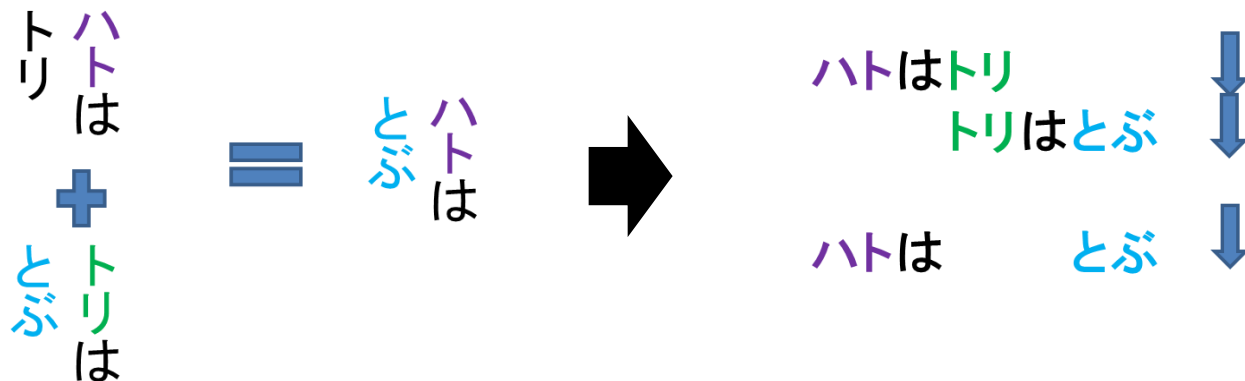
3.2.推論の種類

[ヒト寄りの推論]

- 1) 演繹推論
- 2) 帰納推論
- 3) アブダクション・仮説推論

[機械の推論]

- 4) 非単調推論
- 5) デフォルト推論
- 6) 常識的推論



- 3段階法のモーダスポーネンスは, 人間寄りだがインパクトが小さい.
- 機械への適用は難しいかもしれない.

3.記号推論の発展と適用の検討

3.3 含意関係認識(RTE)の考察

- 2000年代の主要な国際言語タスク
- 含意関係認識とは2つの文を想定
 - ・1つは前提となる文(Text; T)
 - ・もう1つは仮定となる文(Hypothesis; H)を設定,
 - ・それらの関係が成立するかをYes,No(0, 1)で判定するタスク.

⇒BERTによる含意関係認識は、言葉と言葉の距離によるものでしかない

⇒ニュース記事などは、逐次、命題が真であるか確認する必要がある.

例

T: 川端康成は「雪国」などの作品でノーベル文学賞を受賞した.

H: 川端康成は「雪国」の著者である.

→判定: Yes(含意である).

3.記号推論の発展と適用の検討

3.3 含意関係認識(RTE)の考察

一つ一つ, 命題が真であるか確認する必要がある.

例1

T: 彼は朝、コーヒーを飲まずに出かけた。

H: 彼は朝、ここにいた。

→判定: Yes(含意である)。

例2

T: 彼は朝、コーヒーを飲まずに出かけた。

H: 彼は、コーヒーが嫌いだ。

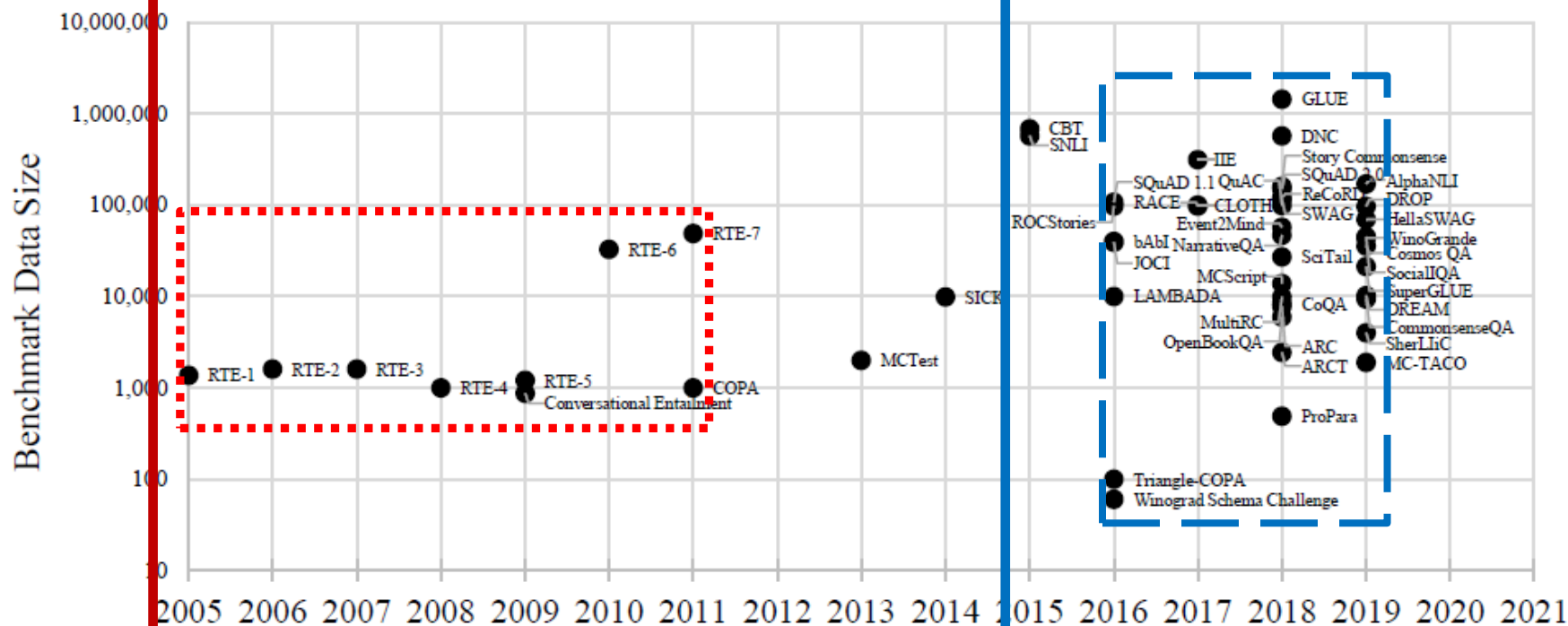
→判定: 不能??

(コーヒーが嫌いだという命題を証明する必要がある)。

3.記号推論の発展と適用の検討

3.3 含意関係認識(RTE)の考察

国際言語タスクはRTEが分岐点となっていた(含意や内包の捉え方).



Wikiの前身となる知識(Cycなど)

深層学習の発展・普及

3.記号推論の発展と適用の検討

3.3 Multi-Hop Reasoningの適用

Multi-Hop Reasoningでの, IF(state)-THEN(action)-BECAUSE(Goal)の論理的アプローチ※の適用を検討する.

- IF節とTHEN節では特定しつつ, BECAUSE節では多様に対応できる.
- BECAUSE節は人の好みによりけりで、対応できる.

IF(state)	THEN(action)	BECAUSE(Goal)
明日, 雨が降るなら,	早く起きよう.	仕事に遅れたくないから.
明日, 雨が降るなら,	早く起きよう.	素晴らしい朝日を見たいから

図4-3-1 IF-THEN-BECAUSEでの捕捉

3.記号推論の発展と適用の検討

3.4 (Multi-Hop Reasoningの)書き言葉への適用

午後8時を過ぎたら、
飲食店での食事は
避けて下さい。

	q	~q
p	pならq「午後8時を過ぎていれば、飲食店で食事できない」	pなら~q「午後8時を過ぎていれば、飲食店で食事できる」
~p	~pならq「午後8時を過ぎていなければ、飲食店で食事できない」	~pなら~q「午後8時を過ぎていなければ、飲食店で食事できる」

→ヒトの認知機能では、論理学の対偶関係で捉えがち。 $p \Rightarrow q$ から $\neg q \Rightarrow \neg p$ が成立するわけではない。

3.記号推論の発展と適用の検討

3.4 (Multi-Hop Reasoningの)書き言葉への適用

午後8時を過ぎたら、
飲食店での食事は
避けて下さい。

	q	~q
p	pならq「午後8時を過ぎ ていれば、飲食店で食事 できない」	pなら~q「午後8時を 過ぎていれば、飲食店 で食事できる」
~p	~pならq「午後8時を過 ぎていなければ、飲食店 で食事できない」	~pなら~q「午後8時を 過ぎていなければ、飲 食店で食事できる」

※1 対偶での $p \Rightarrow q$ から $\neg q \Rightarrow \neg p$ が成立する際、直観主義論理の立場で、排中律の除去を行う必要がある。

※2 排中律とは、「 $p \vee \neg p$ 」の姿勢をとる。

※3 これは含意関係認識のときに特に注意が必要となる。

3.記号推論の発展と適用の検討

3.5 (Multi-Hop Reasoningの)書き言葉への適用

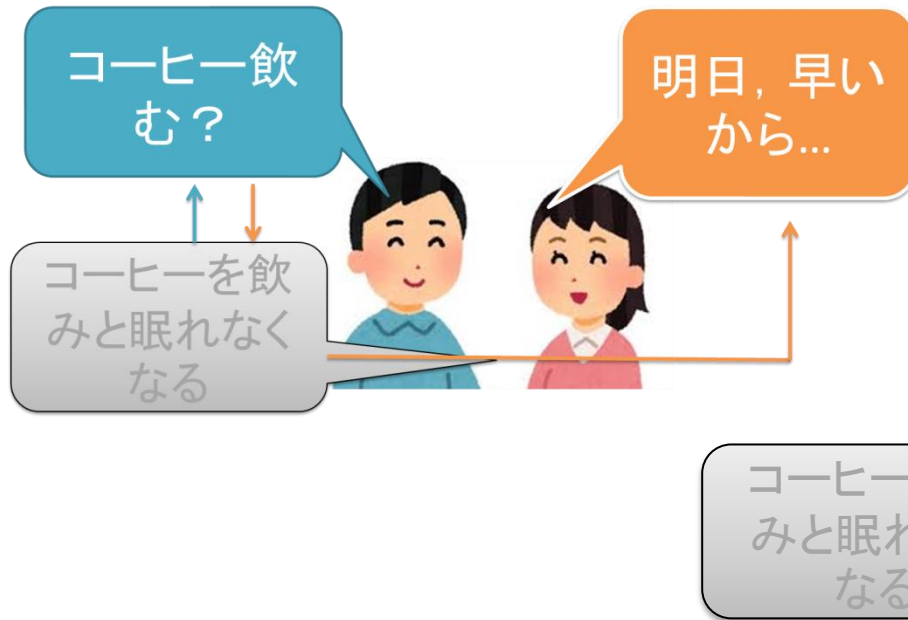
午後8時を過ぎたら、
飲食店での食事は
避けて下さい。

IF(state)	THEN(action)	BECAUSE(Goal)
午後8時を過ぎているなら、	飲食店で食事はできない。	コロナに感染するといけないから。
図4-4-3 IF-THEN-BECAUSEでの捕捉		

→論理学の対偶関係のように危惧する必然はなく, IF-THEN-BECAUSEなら, 字義的にも, そのまま捉えることができる。

3.記号推論の発展と適用の検討

3.6 (Multi-Hop Reasoningの)話し言葉への適用



A(妻)	コーヒー飲む?
B(夫)	明日、早いから...
図4-5-1 夫婦間の会話例	

「コーヒーを飲むと眠れなくなる」

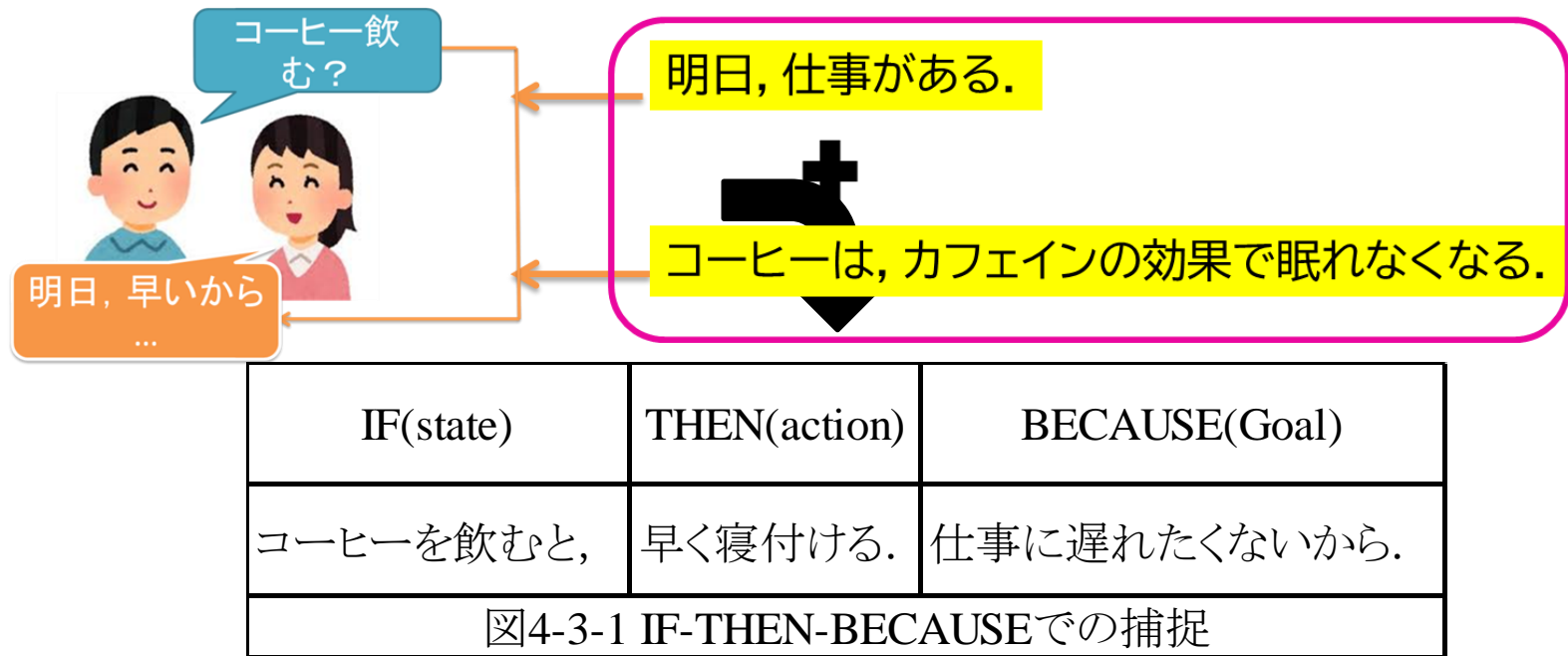
・ 2000年代の課題:

→含意となる本音を内包するものとして捉えるべきか?

→直観主義論理学のように1つ1つ, 事実関係を確認するべきか?

3.記号推論の発展と適用の検討

3.6 (Multi-Hop Reasoningの)話し言葉への適用



- ・「コーヒーにはカフェイン効果があり、寝つきが悪くなることもある」というバラツキがある。
- ・しかし、IF-THEN-BECAUSEなら、「コーヒーを飲むと、カフェイン効果で寝つきが悪くなる」という知識や経験の直接的な引用は不必要になる。

3.記号推論の発展と適用の検討

3.7 (Multi-Hop Reasoningの)関連性理論への適用

娘(真理)	「お母さん、私が明日泊まりにゆくの知ってるの？」			
母(美佐)	「ふとんを干してたわ」			

・関連性理論とは, ポール・グライス(Paul Grice)の4つの公理のうちの「関連性の公理」を発展させた理論.

※「関連性の公理」とは関係のあることをいうべしというもの.
⇒公理に反した場合は, 皮肉などになり, 本音を伝達するという捉え方ができる.

背景	
1)	美佐の母親が今日自宅のベランダにふとんを干していた.
2)	美佐は美佐の母親が今日自宅のベランダにふとんを干していたと言った.
3)	ふとんを干すのは, 宿泊する客が来る前の準備である.
4)	美佐の母親は真理のためにふとんを干していた.
5)	美佐の母親は真理が明日泊まりに来るのを知っている.

3.記号推論の発展と適用の検討

3.8 (Multi-Hop Reasoningの)関連性理論への適用

娘	「お母さん、私が明日泊まりにゆくの知ってるの？」
母	「ふとんを干してたわ」

1)	美佐の母親が今日自宅のベランダにふとんを干していた。
2)	美佐は美佐の母親が今日自宅のベランダにふとんを干していたと言った。
3)	ふとんを干すのは、宿泊する客が来る前の準備である。
4)	美佐の母親は真理のためにふとんを干していた。
5)	美佐の母親は真理が明日泊まりに来るのを知っている。

IF(state)	THEN(action)	BECAUSE(Goal)
明日泊まりに行 くなら、	美佐のお母さんがふとんを干さ ないといけない。	お友達と遊べるよ うに。

・右側の状況(場面=場所+相手+状況)をBECAUSE節に当てはめたい。→人間的な解釈では、やや無理があるかも。

4.実務における常識の定義と課題の考察

4.1南カリフォルニア大学の指摘

BARTやT5で生成された文には、ヒトの常識では有り得ない。

GPT2	A dog throws a frisbee at a football player.
UniLM	Two dogs are throwing frisbees at each other.
BART	A dog throws a frisbee and a dog catches it.
T5	Dog catches a frisbee and throws it to a dog.
図5-1-1 犬がフリスビーを投げる文生成と常識の指摘	

- ・ヒトの常識などの経験則は、論理的なフレームワークを設けるアプローチ。
- ・固有名詞と動詞のセットで、「ヒトが投げて、犬が受け取る」といった論理的な一貫性がある。しかし、機械はそうではなかった。

4.実務における常識の定義と課題の考察

4.2 DARPAの常識の定義とその方向性



DARPA定義:

Machines Common Sense(MCS)

機械における常識に,「人間が周囲の世界と,どのように相互作用するかについての重要な基盤を形成するもの」

→DARPAでは人間の幼少期の発達段階全般を認知能力にマッピングするアプローチ

→幼少期の言語発達はそれほど, 解明にされていない.

→本稿は, 成人した段階を想定.

4.実務における常識の定義と課題の考察

(参考)昨今の常識推論の定義とその方向性

⇒知識獲得から経験や常識推論へシフトしているのが読み取れる。

中島のcommonsenseの定義

・常識推論とは、情報が欠落していても通常は正しい結果を導くことをいう。



Yejin Choiのcommonsenseの定義※2



DEFENSE ADVANCED
RESEARCH PROJECTS AGENCY

- ・常識の定義
- ・実践的な知識と推論の基本レベル
- ・日常の状況やイベントについて
- ・それはほとんどの人の間で一般的に共有されています。たとえば、クローゼットのドアを開いたままにしておくことはできますが、冷蔵庫のドアを開いたままにしておくことはできません。中の食べ物が悪くなる可能性があるからです。

※1「AIがコモンセンスをもつ日」中島 2018 人工知能学会誌 P345

※2 <https://homes.cs.washington.edu/~msap/acl2020-commonsense/slides/01%20-%20Intro.pdf>

4.実務における常識の定義と課題の考察

(参考)昨今の常識推論の定義とその方向性

⇒Choi/DARPAの常識の定義: 賞味期限のイメージと経験や常識(右下の赤枠)

Yejin Choiのcommonsense reasoningの定義※2

- ・常識の定義
- ・実践的な知識と推論の基本レベル

・日常の状況やイベントについて
・それはほとんどの人の間で一般的に共有されています。たとえば、クローゼットのドアを開いたままにしておくことはできますが、冷蔵庫のドアを開いたままにしておくことはできません。中の食べ物が悪くなる可能性があるからです。



体験(たいけん)・行動(こうどう)する

けんこう きつ (03) 健康に気を付ける

■ 活動 - 食べられますか

かい わ れい
(1) 会話例

①

まりあ まりあ: この漬物は、まだ 食べられますか。

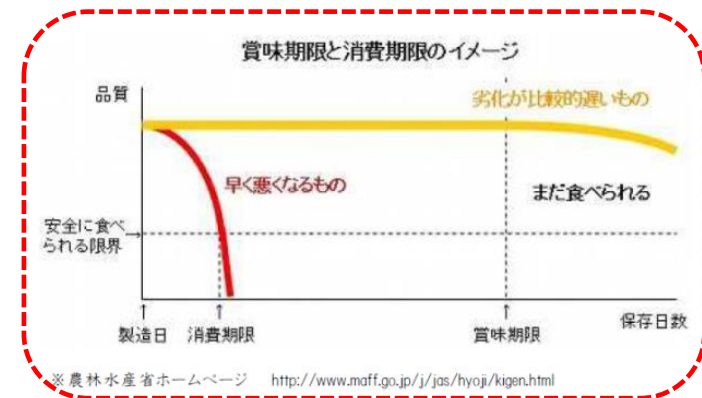
よし子 よし子: 賞味期限は 昨日ですが、大丈夫 でしょう。

まりあ まりあ: そうですか。

②

よし子 よし子: 消費期限ではなく 賞味期限だから、

大丈夫 でしょう。



※「生活者としての外国人」に対する日本語教育の標準的なカリキュラム案 教材例集」

平成24年1月31日 文化審議会国語分科会 P45

4.実務における常識の定義と課題の考察

4.3最高裁判事の常識の捉え方

- ・イギリスの最高裁判事のホフマン卿が、一般常識例を、示したものの(通称, The Starsin事件).

→具体的な貿易取引に従事している者は、他の関係者の業務も知りうるから、過失等が生まれた場合はその責任があるとするもの.

→一般常識には、背景も含まれると見なした.



5. まとめ

1) 第二外国語教育での場面分類を取り上げ, 言語教育と社会文化の区分を考察した.

→ 場所と相手と状況・動機の3つを基本として, 頻繁に出現する順から生活に必要な知識と捉えた.

2) 記号論理学を用いて書き言葉と話し言葉とに適用し考察した.

⇒ Multi-Hop Reasoningの「IF-THEN-BECAUSE」で考察.

⇒ 「IF would」なども含めると, 35%以上の場面での書き言葉・話し言葉がその対象となり得ると考えられる.

⇒ ソーシャルゲームなど, ストーリーがある場面で, 論理推論で捕捉できる範囲は多くあると思われる.

ご清聴ありがとうございました

