

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-

How to approach knowledge acquisition and commonsense reasoning in natural language

-Immigration policy and Japanese language education, from a legal point of view-

太田 博三^{*1}

Hiromitsu Ota

^{*1} 放送大学 教養学部 / (株) VSN

The Open University of Japan/ VSN, Inc. #1

Inference that takes context into consideration, such as the connection between sentences, is required. Sentences made with GPT-3 etc. are as close as possible to those written by humans, but fake news etc. can be understood when it is judged that it is an event that is impossible in commonsense. For literal natural language, supplemented with a reasoning method that can correspond to the machine such as Muti-Hop Reasoning, focusing on the spread range of the viewpoint of language education of domestic and foreign immigration policy and how to grasp common sense in legal judgment. I considered whether it could be done.

1. はじめに

最近、AI ブームが定着し、T5 や GPT-3 などによるフェイクニュースなどが、人と機械との区別がつきずらくなってきている。機械分野における知識量では、Web の情報が圧倒的となり変化した、コミュニケーションの面では十分ではない。ここには、知識を運用するにあたっての経験や常識、暗黙知が関与しているからと考えられる。

自然言語処理の文生成では、文脈を読み取れるようになって来たことが人間に近づいたと言われる大きな要因である。しかし、ニュース記事などの書き言葉であり、対話システムなどの話し言葉となると、十分ではないと言われている。

本稿では、まず、生活を営む上で必要な日本語教育に向けて、人間ベースでの知識獲得やその運用における常識を再考察するものである。次に、機械ベースに適用する際に、どのようなアプローチがより近いかを検討するものである。

2000 年代にグローバル化が進み、国際的な移住が容易になった。日本国内でも、日本語を母語としない海外からの移住者が一定の水準で増加し定着している。ここで、第二外国語の習得が高まれば十分という訳ではない。

海外では、外国人労働者や難民の受け入れに伴って、社会統合として試験が課せられている。言語教育だけではなく、その国の社会や文化の理解も位置づけている。

日本国内では、公的なものではなく、地方自治体やボランティア団体で行われている。自治体では、外国人住民の相談事項とその件数をもとに分類し、多言語表示などで情報が行き届くようにしている。また町内会の回覧板などにも、掲示しアクセスしやすいようになっている。

いずれも、生活に必要な情報を得て、利用できるところまでが求められていると考えられ、知識獲得やその運用による常識が内包され、場面別・人間関係や話者数で分類するものが一般的である。

語用論や認知科学の既存研究では、含意や推意など、対話場面での一部分を取り出して考察する研究が盛んである。これ

ら部分的なものを帰納的に再考察したものが、心の理論や場の理論などである。

さらにポライテネスやメタファー、暗黙知があるが、必ずしも字義的に捉えてわかるものではない表現形式である。生活を送る上で必要不可欠とまではゆかない。友人同士や親子など人間関係の構築がなされているのが前提となり、共通する背景も会話で省略され得る主語や目的語なども、できる限り省略され、そして相手の判断に委ねられる。それらが部分的に・瞬間的に捉える際に、解釈の多義性を楽しむものである。よって、微少な意味の相違はあるものの、生活を送る初段階では、あまり出くわすことも少なく、仮にあるとしても、それほど障害にはならない場面に用いられる。母語話者でも、その解釈は多様で、文脈やその人間関係や会話の背景などを知らないと、なかなか一意に定まらない点で、初心者や機械にとっては難しさもある。

生活のための言語行為の分類は、育児・健康・教育、就労、社会参加、起業など、7-10 種類で、場所と相手と状況とで、位置づけることから始まる。

この分類から、知識や常識を記号論理学の面から考察する。

2. 研究の背景と目的

以下に研究の背景、目的、新規性について記す。

2.1 研究の背景

海外からの移住者に向けた第二外国語習得のためのカリキュラムやオリエンテーションをもとに、場面と相手との関係から想定される分類を検討する。

ドイツやオランダなど欧州諸国では、移民受入れの際に、言語能力とは別に、日常生活を送るための文化や社会を知るオリエンテーションが実施されている。公的な位置づけのもと、専門家が主導となり、行われている。

一方、日本では、自治体の相談窓口での事項と件数から、Q&A として優先順位が決められ掲載されている。言語教育は群馬県大泉市やボトムアップの草の根ルートのようなボランティアによるものが大半で、自治体との連携の上で発展してきた。

ドイツやオランダの言語教育の時間が 600 時間程度(学習者の言語学習の背景等で 400~600 時間)である。

連絡先: 9924658973@campus.ouj.ac.jp

日本では 30 時間程度であり、十分なフォローアップが出来ていない。そこで、文化審議会国語分科会では、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が平成 22 年に提案された。このカリキュラム案(図 2-1-1 参照)は、生活上の基盤を形成する上で必要不可欠であると考えられる生活上の行為の事例であり、大分類・中分類・小分類から構成されている(表 2-1-1 参照)。

標準的なカリキュラム案で扱う生活上の行為(全30 単位)			
○ 健康・安全に暮らす (7 単位)			
・健康を保つ			
・安全を守る			
○ 住居を確保・維持する (2 単位)			
・住居を確保する			
・住環境を整える			
○ 消費活動を行う (4.5 単位)			
・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 (3.5 単位)			
・公共交通機関を利用する			
・自力で移動する			
○ 人とかわる (2.5 単位)			
・他者との関係を円滑にする			
○ 社会の一員となる (4.5 単位)			
・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 (2 単位)			
・余暇を楽しむ			
○ 情報を収集・発信する (4 単位)			
・通信する			
・マスメディアを利用する			
引用先:「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について 文化審議会国語分科会 平成22年5月19日			

図 2-1-1 標準なカリキュラム案

表 2-1-1 大分類・中分類・小分類

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	
・医者の診察を受ける	
・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	

ある生活上の行為が行われるようになるということは、その生活上の行為が行われる場面に関係する言語・社会・文化的な知識を使って行動できるようになるということである。

各地域の自治体やボランティアに適合すべく、ローカライズする必要がある。

これらは、欧州諸国のカリキュラムを踏まえたものでもあり、この分類は場所と相手とを併せてゆくことで、典型的な会話が決まるものと仮定する。

さらに、ボライトネス・関連性理論や皮肉・アイロニーなどでは、前提が異なり、これは何度も顔を合わせている者同士の会話の中でも、出てくるものであるが、場面の分類では、初対面の男女や日本人母語とベトナム人との初対面/女性同士の雑談など、場面分類での示唆が得られる面で、秀逸である。一般的な日本語の教科書では後ろの方に位置づけられているが、ノーマルでの話し言葉では重要である。

2.2 研究の目的

本研究では、人間での知識や常識の場面別の分類を、人工知能を用いた対話システムなどへの適用を検討する。ここでは、記号論理学の帰納法やアブダクション、Multi-Hop Reasoning などで、どのくらい捕捉できるかを考察するものである。

2.3 研究の新規性

研究の新規性に向かうものとして、先行研究では、機械での実装は、ルールベースで医師の診断など専門的なものが多かった。日常生活をなどに必要な知識や常識は、Cyc プロジェクト以来、閉世界での捉え方が難しいといった結論が多く、含意関係認識で終始している。これ以来、深層学習が普及し始めると、推意や事実、さらに常識を仮定の下で内包して捉える記号論理学

小分類				
事例 1（上位項目）				
事例 2（下位項目）		能力記述		場面
		場所	相手	状況・動機
(01) ● 医療機関で治療を受ける				
0101 適切な医療機関の選択をする				
0101060 ★ 隣人に容態を伝えて助言を求める	・隣人に症状を伝えることができる	自宅	隣人	病院へ行きたいが、どの病院に行ったらいいかわからない
	・隣人に適切な医療機関の助言を求めることができる			
	・隣人に適切な医療機関の所在を聞くことができる			

図 2-1-2 小分類と事例と場面

による推論が挙げられるが、果たして、この内包するアプローチでよいのかについても言及してみたい。

まずは、移民など海外からの移住者にとって、生活上必要な知識や常識を場面別に考察し、次に、記号論理学でどのくらいカバーできるかを検討する。最後に、最高裁判事の判決では、常識を、関連性のある場合は相手の立場で予見でき、付随するものとして、捉えるべきと示された。このヒトのアプローチを機械に適用した場合はどのようなかを考察する。

3. 国内外の移住者と言語教育

まず、欧米諸国での海外からの移住者の受け入れに際して、憲法で定められており、特にドイツやオランダなどは、言語教育だけでなく、社会や市民統合に向けて、社会や文化の理解も含めた包括的な教育が義務付けられている。ここでのカリキュラムや制度を考察する。次に、日本での海外からの移住者に対して、地方の各自治体と近隣住民・ボランティアなどが協力して、支援してきた。ボトムアップのアプローチを取り上げる。欧州に比べて公的なものではないため、言語教育の時間も、ドイツやオランダに比べて、10 分の 1 以下であり、包括的なカリキュラムが求められていた。その中で、文化庁が標準的なカリキュラムを提案した。今後の課題としては、前述のトップダウンを自治体レベルで受け止め、再度、ローカライズし、地域社会の慣習などを取り入れることができると、外国人に向けた包括的な日本語教育がカリキュラム上は成立する。時間や費用面はまだ先が見えていないのも現状である。

3.1 欧州諸国の移民政策と言語教育

オランダでは、次の 3 つの段階で、外国人の受け入れの流れを捉えることができる。

- 1) 1945-50 年代の旧植民地からのインドネシア系オランダ人
- 2) 1960 年代の地中海沿岸地域からの労働移民

領域・場面数	場面
市民生活 全10 場面	1)市役所等(個人情報変更の届け出、2)書類申請や各種手続き、3)警察への届け出、4)支払い(銀行)、5)保険、6)住居(住居を借りる、7)公共料金・電話、8)ゴミ・環境、9)教育、10)隣人関係
子育てく 育児、健康、教 育> 全11 場面	1)乳幼児健診センター、2)プレイルーム、3)小学校へ、4)小学校との連絡、5)安全、6)読書と遊び、7)自由時間、8)中等教育へ、9)将来についての話、10)家庭医、11)歯科医
就労分野(一般) 全9 場面	1)職探し(仕事を探す、2)求人応募、3)労働契約についての会話)、4)職場で(労働条件についての会話、5)人事考課面談、6)病欠及び復帰の連絡、7)仕事の打ち合わせ/チームミーティング、8)同僚との話し合い、9)同僚との会話)
就労分野(職種別) 3 種、各5 場面	技術系(顧客とのコンタクト、報告する、労働安全衛生規則への対処、苦情への対応、作業指示の理解)、商業及びサービス業系(顧客とのコンタクト、報告する、労働安全衛生規則への対処、苦情への対応、作業指示の理解)、保健医療及び福祉系(利用者との接触、報告する、労働安全衛生規則への対処、苦情への対応、作業指示の理解)

図 3-1-1 生活場面に必要不可欠なオランダ語

- 3) 1970 年代の旧カリブ植民地のスリナムの移民

また、EU 市民としての滞在や就労しているケースも多くある。これらに対応すべく、生活に必要な不可欠なオランダ語が全部で 45 場面から構成されている(図 3-1-1 を参照)。

また、ドイツでも、1989 年のベルリンの壁の崩壊後、東欧諸国との関係強化から移民を受け入れるようになった。主にトルコや

ユーゴスラビア、そしてポーランドからの移民であり、統合コースのプログラムが用意されている。ここで、ドイツの新移民法は、滞在法と EU 市民の移住の自由に関する法律 (EU 自由移住法) の下で行われている。全移民の 30% が EU 市民であり、移動の自由が認められている。日常生活に困らないドイツ語 (600 時間) とは別に、社会制度や歴史・文化などの理解のためのオリエンテーション (30 時間) が用意されているのが特徴である。日中行われ、受講者の大変は 3-4 ヶ月で終了している。

ドイツ語の能力は、共通欧州準拠枠組 (GER) の能力レベルの 6 段階のうち、3 段階目が求められている。要点を理解でき、熟知したテーマや個人的な関心分野に対して、理路整然と述べることでできると定義されている。ここから、日常生活に必要な語学力だけでなく、やや専門的となる就労先での労働が前提とされていることが読み取れる。たしかに語学教育の時間は 300 時間と多いが、難易度は生活だけでなく、就労も含まれているため、知識の獲得とそれらの利用が求められており、一般的な常識を有していることが期待されていると思われる。

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)	

図 3-1-2 ドイツ語教育及びオリエンテーションのテーマ

3.2 日本の言語教育と標準カリキュラム

日本では、日本語を母語 (第一言語) としない海外からの移住者が来日した場合、地域住民との摩擦が生じないように共生できる環境を形成してゆくために、雇用する側 (企業) と自治体とで協力し合いことから始まった。自治体での相談窓口での内容と件数 (図 3-2-1 を参照) をもとに、日常生活以外で、住居や手続きなどの相談が多く見受けられる。

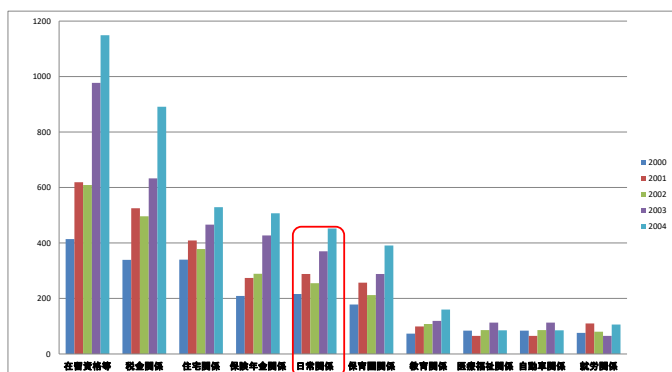


図 3-2-1 群馬県太田市における外国人相談の内容と件数 (2000 年から 2004 年)

当分の間、地元のボランティア活動による日本語教育やサンバカーニバルなど文化的交流も行われてきたが、2008 年から、文化審議会国語分科会で生活に必要な日本語教育の標準カリキュラムが整備され始めた。

ここでの分類は、自治体ベースでの相談内容や件数を参考にしながら、欧州各国の移民政策と言語教育のカリキュラムも考慮されたと思われる。

この大分類を見ると、生活上の行為としての位置づけとその学習順序では、Ⅶ項目目に、他者との関係を円滑にするとして、話し言葉として、対話が挙げられている。日常生活を送る上では、対話はそれほど一番先に来るものではない。健康で安全に暮らし、住環境の整備からものの購入や移動が優先していると考えられる。Ⅶ項目目以降は、地域社会の催しに参加し、余暇を楽しみ、インターネットや電話をする行為がそれ以降に挙げられている。

学習上の分類項目の優先順位では、健康・安全や居住環境、そして物の購入の次に、私的な場面

大分類	小分類	役割
I 健康・安全に暮らす	01 健康を保つ	読み・書きが中心
	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	対話
	06 お金を管理する	
IV 目的地に移動する		対話
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	相手とのインタラクション・相互行為
	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 標準のカリキュラムと学習順序

3.3 ポライテネス・皮肉・アイロニーにおける分類

3.1 や 3.2 では、来日したばかりや対話の初歩に当たる行為といえるが、本項では、比較的、長期間の関係のある人と人との対話で生じるものである。

ポライテネスは相手を敬う上で、丁寧語を使わずに配慮などを表すものもある。また、皮肉は、現実とは反対のことを意図的に言うことで、現実に対する話し手の否定的な態度を暗黙的に (言外に) 伝える言語表現である。

これらを理解するのは、文化的背景からの言語表現の装飾もあるが、経験則から、わかることもあるため、取り上げる (表 3-2-1)。

まず、皮肉・アイロニーを取り上げる。

表 3-2-1 嫁と姑の間の会話断片	
A: 姑	「この料理、おいしいわよね」
B: 嫁	「そっ、そうですか・・・すみません」

次に、ポライテネスの例を取り上げる (表 3-2-1 参照)。ここで、相手との関係と状況・動機の記述を加えると、明確になる。

表 3-2-1 A と B の会話断片 (敬語のない配慮表現)	
A	急いでいる?
B	別に。
A	悪いけど、ペン貸してくれる?
B	いいよ。

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。

4. 記号論理学での捕捉の試み

本節では、3 節でのヒトの生活に必要な日本語を機械でどのくらい捕捉できるかを検討する。まず、記号論理学による表現と推論の視点から入り、次に、含意関係認識(Recognition of Textual Entailment; RTE)を取り上げたい。Multi-Hop Reasoning で会話断片に照らし合わせて、考察する。[Arabishahi 2021]の研究では、人間の 65%に対して、35%であり、SOTA であるとされている。そして、これらを用いて、書き言葉と話し言葉への適用例とを、それぞれ検討する。前者では、論理学の対偶関係を、後者では、省略の補完と含意関係認識との比較を行うものである。

4.1 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として、「IF-THEN」形式のエキスパートシステムが挙げられる。1つ1つの論理命題の列挙が分岐する内容で捉えられるものに限る。日常生活で出てくるものには弱く、医師の診察など専門性の高いものに限られる。この要因の一つに、記号論理学の一階述語論理での表現(図 2.1)にある。この場合、数字の単位は表せず、非常に曖昧なものに限られる。

ここで 1 つ 1 つの知識抽出は Web で可能になったものの、それを捕捉するフレームワークが記号論理学では曖昧なものになり、具体性のあるものは、その対象にならなかった。

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	$\forall x(\text{writer} \rightarrow \exists y(\text{like}(x,y) \wedge \text{documents}))$

図 4-1-1 一階述語論理式の例

4.2 含意関係認識による対象範囲

含意関係認識とは 2 つの文を想定し、1 つは前提となる文(Text; T)を、もう 1 つは仮定となる文(Hypothesis; H)を設定し、それらの関係が成立するかを Yes, No(0, 1)で判定するタスクである。例を以下に示す(図 4-2-1 参照)。

T:	私は昨日、京都で晩御飯を食べた。
H:	私は昨日、京都にいた。
→判定:	含意である

図 4-2-1 含意関係認識の例

ここでは、京都で食事をしていて滞在していた行為とを、直接的に確認する手間を省くものであるが、論理命題となる 1 つ 1 つが確認される訳ではないため、ニュース記事などでは、事実関係が必要になることもあり、注意が必要である。使用される場合は、ボライトネスなどのように、長年、付き合いのある友人や家族同士の会話で、省略する方がごくごく自然な場合が妥当と考えられる。

2000 年代の前半の国際的な言語タスクは、この含意関係認識が主流であった。その後、2019 年頃から BERT による深層学習が出てくると、店舗の広告文と相性の良い消費者の SNS との相性分析が行われるなど、文脈を考慮した活用が消費者の意思を表すとして、マーケティング分野で広く使われるようになった。このように、気持ちや思考などを捉えるのには事実確認がされにくい分野では、利便性は高くなるが、知識や事実関係に照らし合わせる際は、不得手となる。

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あり、論理的推論も加わることが期待されている。

4.3 Multi-Hop Reasoning での適用

Multi-Hop Reasoning には、広い概念のため、本項では、[Arbashahi 2020, 2021] の IF(State)-THEN(action)-BECAUSE(Goal)のアプローチを検討する。深層学習に論理テンプレートを採用したものである。この IF-THEN-BECAUSE は 3 段階法を始めとしたモーダスポーネンスなどの単調推論よりも論理展開が早く、且つ仮設を基にしたアブダクションの推論法を応用したものである。広く事象をカバーすることができるのが特徴である。BECAUSE 節は曖昧に捕捉でき、IF-THEN では統一できるのがメリットである。反実仮定の IF-THEN-, -would-も同じような類型と捉えられる。

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

4.4 書き言葉への適用

最近、よくニュースでも取り上げられたコロナ下の政府の依頼文を取り上げて考察する(図 4-4-1 参照)。ヒトの認知機能では、論理学の対偶関係で捉える傾向があり、ここを起点として考えるが、「午後 8 時前なら、飲食店での食事は構わない」という文脈で捉えてはいけないという政府側の指摘も受け、国民的には混乱した覚えがある。この対偶での解釈では、私たちがよく考えがちで、語用論的には誘導推論と言われている。

午後 8 時を過ぎたら、飲食店での食事は避けて下さい。

図 4-4-1 書き言葉の例: 緊急事態宣言での政府の依頼文

この対偶関係では、論理学上は $p \Rightarrow q$ から $\neg q \Rightarrow \neg p$ が成立するが注意が必要である。直観主義論理の立場で、排中律の除去を行う必要がある。排中律とは、「 $p \vee \neg p$ 」の姿勢をとり、1 つ 1 つの命題を確認するものである。含意関係認識の際も同様の注意が必要となる。

	q	$\neg q$
p	p なら q 「午後8時を過ぎていると、飲食店で食事できない」	p なら $\neg q$ 「午後8時を過ぎていると、飲食店で食事できる」
$\neg p$	$\neg p$ なら q 「午後8時を過ぎていると、飲食店で食事できない」	$\neg p$ なら $\neg q$ 「午後8時を過ぎていると、飲食店で食事できる」

図 4-4-2 対偶関係での解釈

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

IF-THEN-BECAUSE で捉えると図 4-4-3 のようになり、字義的にも、そのまま捉えることができる。

4.5 話し言葉への適用

以下、夫婦間の会話を取り上げて考察する(図 4-5-1 参照)。

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

同じく IF-THEN-BECAUSE で捉えると図 4-5-2 のようになり、捉えることができる。

IF(state)	THEN(action)	BECAUSE(Goal)
コーヒーを飲むと、	早く寝付ける。	仕事に遅れたくないから。

図 4-3-1 IF-THEN-BECAUSE での捕捉

ここでは、「コーヒーを飲むと、早く寝付ける」の間に、「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」という知識が含まれているのか明確ではない。直観主義的には、命題を 1 つ 1 つ、確認する必要がある、ここが知識活用と論理展開との間でネックになっていると思われる。しかし、「コーヒーを飲むと、カフェイン効果で寝つきが悪くなる」という捉え方はやや強引になるため、常識レベルで捉えることも考えられる。

4.6 関係性理論への適用

以下、関連性理論での親と子の会話の背景の捕捉について、考察する。母と娘との会話であり、たった 1 対の対話でしかないが、これを推論するには、単なる文の省略を補完するだけでは不十分であり、母と娘との共通認識や背景の理解が必要となる。背景を5つの場面で記し(図 4-6-1 参照)、会話断片を記す(図 4-6-1 参照)。

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。

図 4-6-1 親と子の会話の背景

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」

図 4-6-2 親と子の会話断片

ここで、IF-THEN-BECAUSE で捉えると、図 4-6-3 のようになり、字義的にも、そのまま捉えることができる。この場面は、文化交流の中でも、ローカル性の高いものだが、よく見かける場面である。

IF(state)	THEN(action)	BECAUSE(Goal)
明日遊びにゆくのなら、	美佐のお母さんがふとんを干さないといけない。	泊まれるように。

図 4-6-3 親と子の会話断片の IF-THEN-BECAUSE での捕捉

5. 最高裁判事の常識と AI の常識

本節では、4.5 での知識と常識との関係をヒトの面と AI など機械の面とで、捉えなおす。以下の図 5-1-1 の因果関係を捉える際の推論の粒度に関係するため、区分しておきたい(図 5-1-1 参照)。A の命題を用いる際は、「コーヒーにはカフェイン効果があり、カフェイン効果には眠れなくなる人もいます」と単調推論で捉えることができる。ナレッジグラフなどでの表現が人によって異なるなどで、画一的にと捉えづらくなる。

一方、A' のように、カフェイン効果を内包する形で、ヒトの個性差を含めて、経験的に常識として捉えると汎用性が高まり、直観主義論理のように 1 つ 1 つの論理命題を検証する必要がなくなる。よって、機械への推論の適用には、常識は 2000 年代の含意関係認識との相性は悪くなかったと言える。

A	「コーヒーを飲むと、カフェイン効果で寝つきが悪くなる」
A'	「コーヒーを飲むと、寝つきが悪くなる」

図 5-1 因果関係の知識活用と推論粒度

ここで、5.1 では南カリフォルニア大学の AI の常識欠如の指摘を取り上げる。5.2 では、DARPA の発達心理学の捉え方を示し、幼児の言語習得のアプローチと本稿のある程度、成人したヒトのアプローチとの比較考察を行い、5.3 でヒトの考える常識を関係性の面から取り上げる。

5.1 最南カリフォルニア大学の指摘と高裁判事の常識

南カリフォルニア大学の Lin や Ren らは、「犬がフリスビーを投げる」という BART や T5 で生成された文に常識が欠けていると指摘し、常識の概念を設けるべきと唱えた。このアプローチは、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を精査するものであり、1 つ 1 つの論理命題を検証することに繋がっていると思われる。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である。

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

5.2 DARPA の常識の定義とその方向性

DARPA では、Machines Common Sense(MCS)として、人間の基本的な背景知識を明確にし、エンコードすることを目的としていると位置づけたうえで、機械における常識を、「ほぼすべての人に共有され、議論を必要とせずに、ほぼすべての人に合理的に期待できるものを認識・理解・判断する基本的な能力」と定義している。また、「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」を常識としており、対話などのインタラクションを対象としていると考えられる。

さらに、コンピュータービジョンなどのオブジェクト(直感的な物理学)や場所(空間ナビゲーション)、そしてエージェント(意図的なアクター)のドメインを含むとしており、言語教育での場面や相手などはこれに適用するものと思われる。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根本的なところからアプローチするものとされているが、本稿は、成熟した段階を想定している面で異なる。

5.3 最高裁判事の常識の捉え方

本節では、イギリスの最高裁判事のホフマン卿は、The Starsin 事件で示した一般的な常識の例を取り上げ、ヒトの常識の捉え方は、関係性の面からであると考えられる。

具体的な貿易取引に従事しているなら、その取引に関与する当事者やその者が属する職能に通じた者につき、それを無理なく入手できるすべての背景的知識を持つ当事者が理解することにより、確定される意味によるべきと、最高裁判事は示したものである。つまり、業務を遂行するにあたり、知識と知識の間に内包する常識が明示され、責任が生じるというものである。

この判決は、船荷証券が国際荷為替による決済で、一般的な常識として、運送契約当事者のみならず、船荷証券の流通にかかわる当事者、とりわけ貿易金融の担保に関係する銀行に共通に理解される知識をもとに、船荷証券の券面記載事項を重視して、文言を解釈すべきことが、宣明された。

ヒトの行為に関係する場面や相手とその行為も含め、常識として知りうるものとして、常識が捉えられている。仕事相手の業務内容や行為がそのスコープとして含まれている。

6. おわりに

本稿では、まず、国内外の海外からの移住者を対象に言語教育だけでなく、その社会文化の理解も含めたカリキュラムを取り上げ、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。その上で、記号論理学の応用である Multi-Hop Reasoning などの「IF-THEN-BECAUSE」など論理フレームで、知識を捕捉できる可能性を確認した。さらに知識から常識へと昇華させる考え方や区別の仕方を検討した。これらが、閉世界の課題に十分ではないかもしれないが、ある状況下であれば、それほど困難とも言えないと思われることが示唆されたと考える。例えば、エージェントベースでのソーシャルゲームなどは、ストーリーが定義できる面で、記述論理学で捕捉できる範囲は少なくない。今後は、アプリなどの実用化に向けて取り組みたい。

参考文献

- [麻生 2020]深層学習と知識・記号推論の融合による AI 基盤技術の発展, 報告書, 研究開発戦略センター (CRDS), 科学技術振興機構 研究開発戦略センター
- [生田 1997] 生田 少子「ポライトネスの理論」大修館書店, 『月刊言語』, pp.66-71
- [井之上 2018]「言語データからの知識獲得と言語処理への応用」人工知能 33 巻 3 号
- [伊東ら 2010]「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について 文化庁 文化審議会国語分科会
- [今野ら 2006]「労働政策研究報告書 No.59 欧州における外国人労働者受入れ制度と社会統合」独立行政法人 労働政策研究・研修機構
- [宇佐美 2020] 宇佐美まゆみ監修「BTSJ 日本語自然会話コーパス(トランスクリプト・音声) 2020 年版」
- [内海 2020]暗黙的提示理論に基づく皮肉理解の認知モデリングへのベイズアプローチ, 第 34 回人工知能学会全国大会 論文集
- [太田 2019] 太田 博三「言語学の語用論や配慮表現の先端技術への適用に関する一考察」第 23 回 SIG-AM 人工知能学会合同研究会
- [河原・野山 2007]『外国人住民への言語サービス』明石書房
- [金田 2008]「生活のための言葉: 国内外先行事例から学ぶこと, 実態調査から明らかにすること」
- [小出 2017]「Common Lisp と人工知能プログラミング(下巻)」オントロノミー合同会社
- [坂原 1985] 坂原 茂「日常言語の推論」, 東京大学出版会
- [椿 2018]「貿易の実務と理論(34)」一般財団法人 日本貿易関係手続簡易化協会
- [時本 2020] 時本 真吾「あいまいな会話はなぜ成立するのか」, 岩波書店
- [増田 2012]「Natural Logic を用いた日本語テキストの含意関係認識」第 26 回 2012 年度人工知能学会全国大会
- [松井 2001]「関連性理論から見たポライトネス」30 巻 No.3 特集 月刊言語 大修館書店
- [米盛 2007]「アブダクション—仮説と発見の論理」勁草書房
- [Arabshahiet.al 2020] "Conversational Neuro-Symbolic Commonsense Reasoning", AAAI Conference on AI
- [Arabshahi et.al 2021]"Conversational Multi-Hop Reasoning with Neural Commonsense Knowledge and Symbolic Logic Rules" arXiv preprint
- [Lin et.al 2020]CommonGen: A Constrained Text Generation Challenge for Generative Commonsense Reasoning, EMNLP2020
- [Mueller 2014] Erik T. Mueller, "Commonsense Reasoning: An Event Calculus Based Approach", Morgan Kaufmann, 2nd edition
- [Norvig 1991] Peter Norvig, "Paradigms of Artificial Intelligence Programming: Case Studies in Common Lisp", Morgan Kaufmann
- [Shwartz 2020] Maarten, Shwartz, Choi et .al., Introductory Tutorial: Commonsense Reasoning for Natural Language Processing, Computational Linguistics, pages 27–33
- [Storks et.al 2019]"Recent Advances in Natural Language Inference: A Survey of Benchmarks, Resources, and Approaches" arXiv