

# 学習者の知識状態を探るための知識依存関係を用いた自動質問生成

川西 哲平<sup>†</sup>兵庫県立大学 情報科学研究科<sup>†</sup>川嶋 宏彰<sup>‡</sup>兵庫県立大学 情報科学研究科<sup>‡</sup>

## 1 はじめに

生徒に教育的な指導を行う方法は大きく一斉指導と個別指導に分けることができる。生徒の理解度や状態を考慮しながら進められる個別指導は一斉指導よりも高い学習効果が得られるが、金銭的、人的なコストが必要であるため、より多くの人へ効果的な個別指導を提供することは難しいのが現状である。

このような背景から、コンピュータを用いた個別指導システムとして Cognitive Tutors [1, 2] が開発された。生徒の行動や知識を表現した認知モデルを用いており、推定した学習進度や理解度を用いて、指導内容に基づくフィードバックや問題解答のヒントを提供できる。しかし学習者が問題解答を行う場面のみを想定しており、想定場面以外で学習者からの質問を受けることや、想定外の質問に対して柔軟な回答を行うことが困難である。一方で、近年急速に発展した大規模言語モデル (LLM) を用いる試み [3] もある。プロンプトでキャラクターやチューターの役割を与えることで、学習支援を行う自然な対話を実現できる。しかし、教育的に正しくない指導内容・順序を含む対話が生成され得るという課題がある。

本研究では、学習理解に必要な知識の依存関係を外部的に用意し、LLM に与えることで、正しい質問を生成する手法を提案する。提案手法の全体図を図 1 に示す。学習理解に必要な知識の依存関係を既存の教材から作成し、生徒からの質問に関連する知識が持つ依存関係を LLM

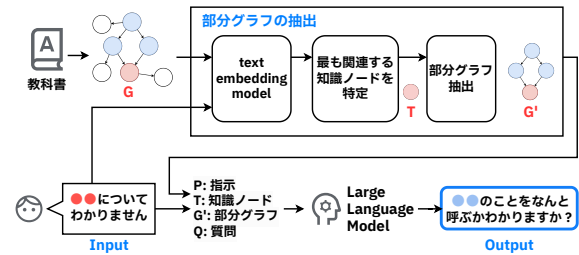


図 1 質問生成システムの全体図

に与えることで、生徒がなぜ理解できていないのかを明らかにするための質問を生成する。

## 2 関連研究

### 2.1 知識ベース型のチュータリングシステム

Cognitive Tutors [1, 2] は、学生が問題を解く際に認知モデルを用いて支援を行うチュータリングシステムである。Cognitive Tutors が生徒に与えるフィードバックやヒントは、問題解答の場面ごとにルールベースで作成されている。そのため、事前に用意されたルールとは異なる行動を生徒が取った場合はエラーメッセージを出力するなど、生徒への働きかけの柔軟性に課題があった。

鈴木ら [4] はチュータリングシステムを構築するにあたり、それぞれの知識の順序を表現する「知識マップ」を e ラーニングの教科書から作成し、フィードバックを行うためのデータベースとして利用している。学習者の発話に含まれる固有名詞が知識マップに含まれるか否かを確認し、フィードバックを生成している。

### 2.2 大規模言語モデルの活用

生徒へのフィードバックやヒントを提示するだけでなく、それらを含む柔軟な対話による学習支援を目指した試み [3] がある。ChatGPT へ入力するプロンプトに、取るべき挙動や指導を行う学習分野、また学習支援の細かな手法を記

Automatic Question Generation Using Knowledge Dependencies to Probe the Learner's Knowledge State

<sup>†</sup> Teppei Kawanishi, University of Hyogo

<sup>‡</sup> Hiroaki Kawashima, University of Hyogo

述することで、柔軟かつ多様な返答による学習支援を実現している。しかし ChatGPT は 2.1 節で取り上げた研究とは異なり、事前に用意された教育的に正しいモデルを所持していないため、誤りのある指導や学習上不自然な順序を含む対話を生成してしまうという課題がある。

そこで本研究では、大規模言語モデルと正しいリソースから作成された知識データとを組み合わせることで、対話に基づく柔軟な学習支援の実現を目指す。

### 3 提案手法

#### 3.1 知識依存関係グラフ

学習者への知識確認のための質問生成を行うためには、教科書などのリソースから作成され、学習理解に必要な知識が学ぶ側にとって自然な順序で記載されているデータが必要である。本研究では知識に関するデータを、知識同士の依存関係に着目して作成した。以下これを**知識依存関係グラフ**と呼び、 $G = (V, E)$ 、 $V = \{v_i | i = 1, \dots, n\}$ 、 $E = \{e_i | i = 1, \dots, n\}$  と定義する。 $v_i$  は「知識ノード」であり、「A は B である」といった説明文（宣言的知識）とする。また  $e_i$  は各ノード間の依存関係を表す有向エッジである。以下では、数学 I の教科書 [5] の第 2 章において、内容を理解するために必要とされている箇所の説明文をノードとし、著者によって依存関係の有無を判断することで作成した、77 ノード、131 エッジの知識依存関係グラフを用いる。

$G$  における知識ノード  $v$  は、knowledgeID, explanation, name の 3 つのプロパティを持つ。knowledgeID は各ノードにつけられたユニークな値である。explanation は教科書から抽出した宣言的知識を表す文章が、そして name には explanation の内容をもとに独自に付与した短い名前が含まれている。

作成した知識依存関係グラフの一部を図 2 に示す。図 2 の例では、「値域とは」と名付けた一番右にある知識ノードを理解するためには、その知識ノードに入るエッジを持つ知識ノード（「定義域とは」と「値とは」）の両方を理解しておく必要があることを表現している。また、「定

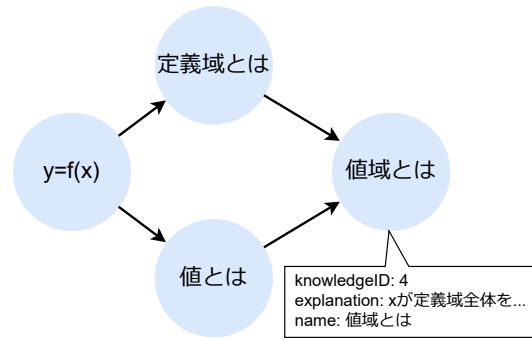


図 2 作成した知識依存関係グラフの一部

義域とは」と「値とは」という 2 つの知識ノードは、さらに左にある知識ノード（「 $y=f(x)$ 」）からのエッジを持っている。このように、「値域とは」という知識の理解には、その知識ノードにつながる 3 つすべての知識を理解しておく必要があることを表現している。

#### 3.2 大規模言語モデルを用いた質問生成

提案する質問生成システムの全体図を図 1 に示す。システムは入力として、学習者からの質問  $Q$  を得る。その後、入力された質問  $Q$  に最も関連している知識ノード  $T$  を、3.1 節で述べた知識依存関係グラフ  $G$  内において特定する。そして、特定した最も関連する知識ノード  $T$  を理解するために必要な部分グラフ  $G'$  を抽出する。ここで  $G'$  は、 $T$  を理解するために必要なすべての知識ノードのみによって構成され、知識ノード  $T$  から芋づる式に抽出される。

次に、部分グラフ  $G'$  をもとに ChatGPT に Context として入力する表形式データを作成する。データは CSV 形式で記述され、各行は 2 つのノードをひとつのペアとし、エッジが出るノードの knowledgeID と explanation, 同様にエッジが入るノードの knowledgeID と explanation を持つ。ソースコード 1 に表形式データの例の一部を示す。

##### ソースコード 1 抽出したデータ

```
index, start knowledgeID, end knowledgeID,
  start explanation, end explanation
0, 0, 3, 2つの整数..., y=f(x)において...
1, 0, 12, 2つの整数..., xの1次式で...
```

最後に、学習者からの質問  $Q$ ,  $Q$  に最も関連

する知識ノード  $T$ , そして ChatGPT への指示文章をプロンプトに含め入力し, 学習者へすべき質問を出力する。

### 3.2.1 部分グラフの抽出

はじめに, 学習者の質問  $Q$  と最も関連する知識ノード  $T$  を知識依存関係グラフ  $G$  内から特定する。ここでは, 学習者からの質問  $Q$  と各知識ノードが持つ explanation を埋め込みモデルを用いてそれぞれベクトル化し, コサイン類似度が最大となるノードを  $T$  とする。その後, 特定した知識ノード  $T$  を起点とし,  $T$  を理解するために必要なすべての知識ノードを部分グラフ  $G'$  として抽出する。

### 3.2.2 ChatGPT による質問生成

タスクや使用すべき情報を ChatGPT に指示するプロンプトである System プロンプトを以下のように設計した。

- Instruction (ロール, 状況付与, タスク指示)
- Context ( $T, G'$  から作成した CSV)
- Output (出力形式の指示)

「Instruction」ではシステムと対話相手がどのような関係であるか, また行ってほしいタスクとそのタスクを達成する方法を指示している。

「Context」では, 学習者からの質問に応じて効果的な質問生成に必要なデータを, 知識依存関係グラフの部分グラフ  $G'$  から作成した CSV 形式のデータとして入力している。学習者からの質問に最も関連するノード  $T$ , および部分グラフ  $G'$  から作成した CSV 形式データには, knowledgeID と explanation の情報を含めている。これは, ChatGPT において, knowledgeID は各知識ノードの関係性を理解するのを容易にし, explanation は生成する質問の情報源として使用できると期待されるためである。

「Output」では生成質問の出力形式を指示している。学習支援の場面において学習者へ質問する際には, 適切な足場架けとなるヒントを提示するために, どの知識を理解できているか確認する必要がある。そのため, 理解の有無に関する回答が得られるような言葉遣いを指示している。さらに不必要な情報の出力を防ぐため, 生成質問文とその質問文のもととなった知

識ノードの情報のみを出力するよう指示している。

## 4 評価実験

学習者を想定した質問を設定するとともに, その質問に最も関連する知識ノードのペアを 8 種用意し, 各質問に対して 5 回の質問生成を行った。

提案手法による評価実験では, GPT-3.5 と GPT-4 をベースにした ChatGPT を用いた。すべての条件で temperature (温度) は同一とし, 多様な出力結果を得るために 1.0 とした。GPT-3.5 と GPT-4 では同一のプロンプトを用い, それぞれ計 40 回の出力を行った。知識依存関係グラフに関する情報を与えずに出力を行った結果をベースライン (LLM-only) として用いた。ベースラインにおける入力プロンプトは, 提案手法における入力プロンプトから知識依存関係グラフに関する部分を削除し, できる限り近い文章とした。

### 4.1 評価指標

各生成質問がどの知識ノードに関する質問であるかに基づいて評価を行った。本実験において生成された質問は, 以下の 4 種類に分類できる。また知識依存関係グラフにおける各種類の対応ノードを図 3 に示す。

- $T$  を理解するために必要な知識ノードについての質問生成
- $T$  自体についての質問生成
- $T$  の理解を前提とした知識ノードについての質問生成
- $T$  に直接関連しない知識ノードについての質問生成

生成質問が A に対応する知識ノードについての質問である場合は, 学習者の知識を探るための質問として適切である。そのため,  $T$  との距離が小さいほど学習者からの質問により関連した質問の生成ができていると考え, 最短距離を評価指標とした。生成質問が B, C, D に対応する知識ノードについての質問である場合は, 学習者の知識を探るための質問として不適切である。そのためそれぞれの場合についての質問

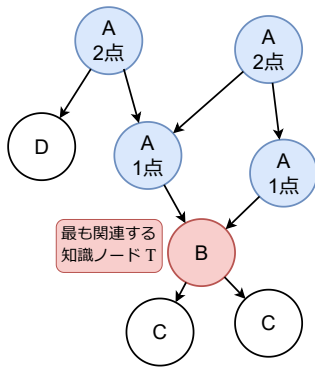


図 3 各生成質問における知識依存関係グラフとの対応

表 1 質問生成の評価

Method (ver.)	dist↓	A↑	B↓	C↓	D↓
LLM-only (3.5)	1.56	9	13	5	0
LLM-only (4)	2.46	26	<b>2</b>	2	5
Proposed (3.5)	2.43	14	17	0	0
Proposed (4)	<b>1.46</b>	<b>33</b>	7	0	0

の数を調べた。

#### 4.2 評価結果

評価を行った結果を表 1 に示す。dist は生成質問が A であった場合の最短距離の平均であり、A, B, C, D はそれぞれのノードに関する質問が観測された数を表す。なお、Method の括弧内の数字は GPT のバージョンを表す。

##### 4.2.1 ベースモデルによる比較

B, C, D の和が最も少なかったのが GPT-4 を用いた提案手法による出力、次に少ないのが GPT-4 のみを用いた出力の場合である。つまり GPT-4 を用いた 2 つの場合には、より安定して知識依存関係に基づく質問を生成できたといえる。一方で GPT-3.5 を用いた場合は、LLM-only と Proposed のそれぞれにおいて、学習者の質問に最も関連する知識ノード  $T$  自体について質問した数 (B) が GPT-4 を用いた場合よりも多く、学習者の知識を探る質問ができていない。このことから、GPT-3.5 と GPT-4 のモデル性能差が生成結果に大きく影響していることがわかる。

##### 4.2.2 知識依存関係グラフの有無による比較

提案手法においては、学習者からの質問に最も関連する知識ノード  $T$  とは関係のない知識に関する質問 (C と D) の生成は観測されな

表 2 学習者からの質問「2 次関数とはなんですか」に対する各手法の生成質問

LLM-only (3.5)	2 次関数のグラフの特徴について知っていますか？
LLM-only (4)	1 次関数とはなんですか、説明できますか？
Proposed (3.5)	2 次関数とはなんですか？
Proposed (4)	$x$ の 1 次式で表される関数をどう呼ぶか知っていますか？

かった。このことから、正しい学習順序・依存関係に基づいた質問生成を行うために知識の依存関係を外部知識として入力することは効果があるといえる。また、GPT-4 を用いた提案手法による出力では最短距離の平均が最も小さいことから、学習者からの質問に最も関連する知識  $T$  との依存関係に基づいた質問生成ができていない。GPT-3.5 のみの場合も平均最短距離は小さいが、知識依存関係に基づかない質問生成数 (B, C, D の和) が 18 回と最多である。

#### 4.3 生成例

本実験において生成された質問の一部を表 2 に示す。GPT-3.5 を用いた 2 つの場合では、学習者からの質問に最も関連する知識ノード  $T$  についての質問 (B) を生成している。一方で GPT-4 を用いた場合は、2 次関数を理解するために必要な知識である 1 次関数についての質問をともに生成している。特に GPT-4 を用いた提案手法の場合には、GPT-4 のみの場合よりも学習支援の場において自然な言葉遣いによる質問を生成できている。

謝辞 本研究の一部は科研費 21H05302 の補助を受けて行った。

#### 参考文献

- [1] John Anderson, Albert Corbett, Kenneth Koedinger, and Ray Pelletier. Cognitive Tutors: Lessons Learned. *Journal of the Learning Sciences*, Vol. 4, pp. 167–207, April 1995.
- [2] Kenneth Koedinger and Albert Corbett. Cognitive Tutors: Technology Bringing Learning Science to the Classroom. January 2006.
- [3] Learn from the Best of the class. <https://home.x-in-y.com/course/view.php?id=157>.
- [4] 鈴木舜也, 吉野孝. 講義における学生の質問行動を促進するチャットボットの評価. *ヒューマンインタフェース学会論文誌*, Vol. 25, No. 2, pp. 77–88, 2023.
- [5] 高校教科書 改訂版 数学 I [数 I / 327]. 数研出版, 2017.