

細分化した学習項目に基づく 数学学習用問題の AI 自動生成に関する検討

石田 崇^{*1}・須子 統太^{*2}・小林 学^{*2}・雲居 玄道^{*3}・隈 裕子^{*4}・平澤 茂一^{*2}

Email: ishida@tcue.ac.jp

*1: 高崎経済大学
*2: 早稲田大学
*3: 長岡技術科学大学
*4: 九州産業大学

◎Key Words 生成 AI, 問題自動生成, 数学学習, 因数分解

1. はじめに

近年、教育分野における AI 技術の活用が急速に進展している。特に、大規模言語モデル (Large Language Models: LLMs) の登場により、教育コンテンツの自動生成や学習支援など多様な用途での応用が期待されている⁽¹⁾⁽²⁾⁽³⁾。これらのモデルは、人間のような自然なテキスト生成能力を持ち、教員や学習者の個別ニーズに応じた教材作成に活用できる可能性がある。一方で、学校教育においては「個別最適な学び」の実現が求められており⁽⁴⁾、学生や生徒一人ひとりの理解度や学習進度に応じた教材を提供することは重要な課題である。

これまでに、生成 AI を活用して学習用の練習問題を自動作成する研究も存在しているが⁽⁵⁾⁽⁶⁾⁽⁷⁾、著者らは統計学における仮説検定を対象とした練習問題の自動生成手法の検討を行った⁽⁷⁾。ここで作成する検定の問題は、単なる数値計算問題ではなく、数値データの背後に実社会とリンクした文脈をもつ文章問題であり、AI が持つ文脈生成の特性を活用したものである。出題したい検定の種類 (上側/下側/両側) や解答のパターン (棄却する/しない) を指定しつつ、適切な文脈をもち数学的に成立する問題文を生成するためのプロンプトの設計し、問題文から解答や解説文を自動作成する作問システムについて検討を行った。

仮説検定の問題作成が AI の文脈生成能力に着目したのに対し、本研究では学習問題を解くために必要な知識やスキルの段階的な階層構造をモデルとして AI に与えて認識させ、この構造に応じた適切な問題を生成する能力に着目するものである。学習を段階的に進める場合には、どのような内容をどのような順序で学ぶかが重要である⁽⁸⁾。段階的な学習において知識の定着を図るためには、それぞれの段階ごとに内容やレベルが統一された多様な多数の練習問題が用意されることが望ましい。

そこで本研究では、学習項目やレベルの階層構造が他に比べて比較的明確と考えられる数学に着目し、特に具体的な事例として因数分解の学習問題を取り上げて検討することにする。因数分解は、共通因数の発見、乗法公式の適用、複合的な手法の組み合わせなど、難易度の階層構造や学習項目間の順序関係が分かりやすい。まず、学習のための階層モデルを与えられたもとで、各段階に応じた問題を適切に生成するためのプロンプトを設計し、実際に AI に生成させた問題群の評価を通じて、本手法の有効性を検証する。

なお、本稿では階層モデルそのものの妥当性や構築方法については議論の対象とはしない。学習項目間の関係性の分析や構造化の手法に関する研究も存在しているが⁽⁸⁾⁽⁹⁾⁽¹⁰⁾⁽¹¹⁾⁽¹²⁾、ここでは学習項目の階層モデルは教員が目的に応じて設定したものが与えられている前提とする。このモデルを用いて、生成 AI により効率的かつ正確に段階ごとの問題を作成できるのかという課題に焦点を当てて、階層別問題生成のためのプロンプトを設計し、実際に生成した問題の品質の評価を行う。

本稿では具体的な事例として因数分解を取り上げて検討するが、本研究の意義は因数分解の問題生成にとどまらず、階層構造を持つ様々な学習領域において、AI を活用した系統的な問題生成の可能性を示すことにある。

2. 階層構造モデルに基づく問題生成

2.1 階層構造モデル

基礎的な概念から発展的な内容へと段階的に学習を進める場合、学習項目は下位から上位レベルの内容へと階層的に構成される。階層的な学習項目が満たすべき一般的な要件として、以下が考えられる。

- 下位レベルの知識・技能が上位レベルの前提となる
- 各レベルで習得すべき内容が明確に定義されている
- レベル間の移行に必要な条件が明示されている

2.2 因数分解におけるスキルモデル

本研究では、上記の枠組みを具体的に検討するための事例として、因数分解を取り上げる。因数分解は階層構造を明確にしやすく、本研究で最初に検討する事例として適した特徴を持つと考えられる。

因数分解における学習項目の階層モデルとして採用したものを表 1 に示す。このモデルでは L1~L4 の 4 つの主要レベルから構成され、もっとも基本的と考えられる共通因子で括る L1 レベルから、L4 へ進むにつれて問題を解くために必要なスキルの難易度が上がる。また 4 つのレベルはさらに小項目に細分化されている。

なお、ここでは階層モデルは最適なモデルが一意に定まるものではなく、問題作成者が出題の意図や目的に応じて設定したものを扱い、このモデルにしたがって問題を AI が生成できるかどうか評価することとする。モデルの是非については必要があれば別途検討・改良し、それに伴ってプロンプトも適宜修正して用いる。

表1: 因数分解の学習階層モデル

レベル	例
L1: 共通因数	
L1-1: 基本的な共通因数 (正で括る)	
L1-1-1: 単一の数値	$3x + 12$
L1-1-2: 単一の文字	$ax + ay$
L1-1-3: 共通素因数	$12x + 18y$
L1-1-4: 複数の文字・次数	$6x^2y - 10xy^2$
L1-2: 負の数で括る	$-2a - 4b$
L1-3: 多項式の共通因数	$a(x + y) + b(x + y)$
L2: 乗法公式 (基本形) (x^2の係数が1)	
L2-1: 2項 (A^2-B^2)	$x^2 - 25$
L2-2: 3項 ($(A \pm B)^2$)	
L2-2-1: $(A+B)^2$	$x^2 + 10x + 25$
L2-2-2: $(A-B)^2$	$x^2 - 10x + 25$
L2-3: 3項 ($(x+a)(x+b)$)	
L2-3-1: ($a, b > 0$)	$x^2 + 7x + 12$
L2-3-2: ($a, b < 0$)	$x^2 - 7x + 12$
L2-3-3: ($a \times b < 0$)	$x^2 + 4x - 12$
L3: 乗法公式 (発展形) (x^2の係数が1でない)	
L3-1: 2項 (A^2-B^2)	$4x^2 - 9y^2$
L3-2: 3項 ($(A \pm B)^2$)	$9x^2 - 12x + 4y^2$
L3-3: たすき掛け	
L3-3-1: (すべて正の項)	$2x^2 + 7xy + 3y^2$
L3-3-2: (負の項を含む)	$3x^2 - 5xy - 2y^2$
L4: 複合テクニック	
L4-1: 共通因数→公式	
L4-1-1: (L1→L2)	$2x^2 - 50$
L4-1-2: (L1→L3)	$6ax^2 + ax - 2a$
L4-2: 置き換え	
L4-2-1: (置き換え→公式)	$(x + y)^2 - 5(x + y) + 6$
L4-2-1: (符号→置き換え)	$a(x - y) - x + y$
L4-3: 項の組み合わせ	
L4-3-1: 4項(2+2)	$ax - bx + ay - by$
L4-3-2: 4項(3+1)	$x^2 + 2x + 1 - y^2$
L4-4: 1つの文字で整理	$x^2 + xy - 3x - y + 2$

2.3 問題生成の要件定義

数学の学習問題を想定し、階層構造を持つ一般的な学習項目に対して問題を生成する際の要件を以下のように定義した。

1. レベル適合性の要件

- 生成される問題は、指定されたレベルの本質的特徴を正確に反映すること。
- 指定されてレベルよりも下位レベルのスキルのみで解ける問題は生成しない (簡単すぎる問題の排除)。
- 上位レベルのスキルを必要とする問題は生成しない (レベル逸脱の防止)。

2. 数学的妥当性の要件

- 生成される式が数学的に正しい問題であること。
- 解答が一意に定まること。
- 計算過程で不合理な数値 (極端に大きい係数など) が現れないこと。

3. 多様性の要件

- 同一レベル内で、パラメータや形式が多様であること。
- 機械的な暗記では対応できない問題セットとなること。
- 実際の教育現場で扱われる範囲を網羅すること。

以上の条件をプロンプトとして記述して生成AIに正確に認識させ、要求する問題を多数作成することが本研究の目指す課題である。この課題は、因数分解に限らず、階層構造を持つあらゆる学習領域に共通する問題である。

3. プロンプトの設計

3.1 基本的なプロンプトの構造

基本的なタスクとして以下の内容をプロンプトに記述した。

#役割設定

「あなたは経験豊富な数学の教師であり、教材作成のエキスパートです。生徒の習熟度に合わせて質の高い学習用練習問題を作成することを得意としています」

#タスク

以下の階層モデルに厳密に従って、因数分解の練習問題を生成します。

#階層モデル定義

各レベルごとの要件

[レベルID][タイトル][レベルの説明][問題例]

#指示

上記のモデルに基づいて、以下の条件に合致する因数分解の問題を生成してください。

- 作成するスキルID
- 作成する問題数
- 出力形式 (例えば、問題と解答の式のみをリスト形式で出力するなど)

#要件

- レベル適合の要件 (略。2.3 節参照)
- 数学的妥当性の要件 (略。2.3 節参照)
- 多様性の要件: 生成する問題群には、以下の要素をバランス良く含めることで、多様性を確保してください。
 - (係数の多様性) 係数が2以上の場合は、1桁の簡単な係数から少し計算が必要な2桁の係数まで幅広く使用する。
 - (負の数の配置) 式の中に含まれる負の数の位置や個数を多様にする。
 - (使用文字の多様性) x や y だけでなく、 a, b, m, n などの異なる文字を使用する。
 - (項の順序) 標準的な降べきの順だけでなく、学習者が自分で式を整理する必要があるように、項の順序を入れ替えた問題も一部含める。
 - (類似パターンの混在) 同じスキルID内でも、解法のポイントが微妙に異なる類似パターンを混在させる。
- 条件・要件の厳守: 生成する問題は、指定されたスキルIDの定義とレベル、指示、要件に厳密に従ってください。

このプロンプトにより、問題作成者が意図する階層モデルの構造と作成の指示を生成AIに対して明確に与えることが可能となった。

3.2 レベル別プロンプトの詳細設計

各レベルの階層モデルについて、その特徴を正確に反映したプロンプトを設計した。以下に代表的な例を示す。

(1) L1 レベルの例

- level_id: "L1"
title: "共通因数"
description: "すべての因数分解の土台。式全体を俯瞰し、共通の要素を見つけ出す「構造的把握能力」を養う。"

sublevels:
- level_id: "L1-1"
title: "単純な共通因数（正の項）"
description: "認知負荷が最も低い導入段階。分配法則の逆演算としてスムーズに理解できる。"

children:
- level_id: "L1-1-1"
title: "単一の数値"
description: "具体的な数値を扱うため、直感的。"
example: "2x+6"
- level_id: "L1-1-2"
title: "単一の文字"
description: "抽象度が少し上がるが、文字式に慣れていれば容易。"
example: "ax+ay"
- level_id: "L1-1-3"
title: "共通素因数"
description: "係数を素因数分解する一段階上の思考が必要になる。"
example: "12x+18y"
- level_id: "L1-1-4"
title: "複数の文字・次数"
description: "複数の文字とそれぞれの指数を正確に比較する必要がある、注意力が求められる。"
example: "6x^2y - 10xy^2"

- level_id: "L1-2"
title: "符号の扱い"
description: "負の数が入ることで、符号のルールを正確に適用する必要が生じる。つまずきが増え始める。"

children:
- level_id: "L1-2-1"
title: "負の項を含む（正で括る）"
description: "括弧の中に負の項が残るだけなので、負荷は比較的小さい。"
example: "3x-9y"
- level_id: "L1-2-2"
title: "負の数で括る"
description: "つまずきポイント。括りだした後の括弧内の項、すべての符号が反転することを見落としやすい。"
example: "-2a-4b"

- level_id: "L1-3"
title: "多項式の共通因数"
description: "具体的な数や文字ではなく、「式の塊」を一つの要素として認識する抽象的な思考への飛躍が求められる。"
example: "a(x+y)+b(x+y)"

(2) L3 レベルの例

- level_id: "L3"
title: "乗法公式（発展形）"
description: "x^2 の係数が 1 でないパターン。係数まで含

めて構造を捉える必要があり、計算の複雑さが格段に上がる。"

sublevels:
- level_id: "L3-1"
title: "2 項 : A^2-B^2"
description: "係数が平方数であることに気づく必要がある。"
example: "4x^2-9y^2"
- level_id: "L3-2"
title: "3 項 : (A ± B)^2"
description: "係数を含めて両端が平方の形になっているかを見抜く、より高度なパターン認識が必要。"
example: "9x^2-12xy+4y^2"
- level_id: "L3-3"
title: "たすき掛け"
description: "最大の壁。体系的な試行錯誤が求められる。係数の約数をリストアップし、組み合わせを検証するプロセスは負荷が高い。"

children:
- level_id: "L3-3-1"
title: "全て正の項"
description: "組み合わせの探索に集中できる導入段階。"
example: "2x^2+7x+3"
- level_id: "L3-3-2"
title: "負の項を含む"
description: "符号の組み合わせが加わることで、探索空間が爆発的に増え、混乱を招きやすい。"
example: "3x^2-5x-2"

4. 評価実験

本研究では、生成 AI のモデルとして Gemini 2.5 Pro を使用し、各スキルレベルに対して 20 問ずつ、計 460 問 (23 カテゴリ×20 問) を実際に生成して問題の評価を行った。

4.1 生成結果の定量的評価

生成された問題について、以下の観点から定量的評価を実施した。

(1) レベル適合率

生成された 460 問を手でチェックし、レベルの不適合がないかを確認した。その結果、今回作成した問題においてレベルの不適合を 1 問確認した（不適合率 0.21%）。

不適合の具体例は以下の通りである。

例. ID : L3-1 [乗法公式（発展形）. x^2 の係数が 1 でない A^2-B^2 の形]
問題 : $100p^2 - 225q^2$ ($25(2p + 3q)(2p - 3q)$)
不適合の理由 : まず共通因数 25 で括ることができてしまうことから、純粋に L3-1 レベルの問題になっておらず、今回想定している階層モデルでは L4-1-2 に分類される方がより適切と考えられる（完全な不適合とは言い切れない）。係数が大きくなってしまったことで意図せず共通因数を含んでしまったためと考えられる。

(2) 数学的妥当性

生成された問題が正しく因数分解可能かを確認した。その結果、今回作成した問題において数学的妥当性の不適合を 2 問確認した（不適合率 0.43%）。不適合の具体例

は以下の通りである。

例 1. ID: L2-3-2 [乗法公式 (基本形). 3 項 $(x+a)(x+b)$, $a, b < 0$]

問題: $c^2 - 16p + 55 \quad ((p-5)(p-11))$

不適合の理由: 変数 c と p が混在してしまった。これは誤植によるもので数学的に解けない問題である。

例 2. ID: L4-2-1 [置き換え→公式]

問題: $(p+q)^2 - (r+s)^2$
 $((p+q+r+s)(p+q-r+s))$

不適合の理由: 提示された解答が間違っているケース。

問題は適合しているので問題自体の不備では無い。

4.2 生成例と質的評価

各レベルで実際に生成された問題の一例を以下に示す。

(1) L1-1-1 (基本的な共通因数) の生成例:

$$6x + 12, 7a - 21, 8m + 8, 17b + 34, 9n - 36, \\ 6x + 18y - 6$$

これらの問題は、指定されたレベルの要件を満たしており、係数の選択も適切である。また、文字や符号、項の並び順も多様化されている。

(2) L2-3-3 (3 項 $(x+a)(x+b)$, ab が負) の生成例:

$$x^2 + 2x - 15, a^2 - 3a - 10, y^2 + 4y - 21, \\ n^2 - 2n - 168, a^2 + a - 380,$$

このレベルでも、適切な問題が生成されていると考えられる。和と積の関係が明確であり、適切な難易度となっている。ただし、一部で積 (ab) の値が大きく和と積の組み合わせを見つけるのが困難な数値も含まれていた。

5. 結果のまとめと考察

(1) 適合率の高さ

生成した 460 問について、99%以上の高い精度でレベルに適合する数学的に正しい問題が作成できた。

(2) レベル境界の曖昧さ

階層の定義に曖昧さがあると 2 つ以上のカテゴリに分類されるケースが生じる。このような境界の例を排除するためにはルールを細かく設定しておくことも必要であるが、曖昧性を完全に無くすことは難しいと考えられるので、最終的に人手によるチェックと判断は必要である。

(3) 同一レベル内での難易度の差

同じカテゴリの同類の問題であっても係数が大きすぎて計算が煩雑になることで難易度に差が生じるようなケースがあった。係数の絶対値の制御なども必要と考えられる。

(4) レベル別問題生成の実現可能性

適切なプロンプトにより高い精度で階層別に問題を作成できることが示されたが、これは LLMs が単に文脈を生成するだけでなく、あたかも数学問題の背後にある構造を理解しながら問題を作り分けることが可能であるかのような振る舞いによるものである。因数分解に限らず、階層構造を持つ他の学習領域にも適用可能であるかの検証が必要である。

(5) 学習問題生成の効率化

最終的に人手によるチェックは必要であるものの、学習者の理解度に応じたきめ細かい段階ごとに多様な問題を大量に作成することが可能となった。

6. おわりに

本研究では、細分化した学習項目に基づく数学学習用問題の AI 自動生成についてプロンプトを設計し、指定された階層構造に応じて高い精度で問題が作成できた。作成する問題の項目や難易度レベルを細かく制御できることで、学習者一人ひとりの到達度に応じた問題提供が技術的に可能となり、個別最適な学びの実現に寄与できると考えられる。

今後の課題として、数学の他の単元 (方程式、関数、図形など) への展開や、他教科への応用、項目間に複雑な関係性を持ち、深い思考を必要とするより高次の学習内容における問題作成への応用などが挙げられる。

参考文献

- (1) Peláez-Sánchez IC, Velarde-Camaqui D and Glasserman-Morales LD: “The impact of large language models on higher education: exploring the connection between AI and Education 4.0”, *Frontiers in Education* Vol.9, doi:10.3389/feduc.2024.1392091 (2024).
- (2) Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D.: “Practical and ethical challenges of large language models in education: A systematic scoping review”, *British Journal of Educational Technology*, Vol.55, No.1, pp.90-112, doi:10.1111/bjet.13370 (2024).
- (3) 原田隆史: “大学教育現場における生成 AI 技術の利用”, 情報の科学と技術, 74(8), pp. 298-303, 情報科学技術協会 (2024).
- (4) 文部科学省: “「令和の日本型学校教育」の構築を目指して～全ての子供たちの可能性を引き出す、個別最適な学びと、協働的な学びの実現～ (答申) (中教審第 228 号)”, https://www.mext.go.jp/content/20210126-mxt_syoto02-000012321_2-4.pdf, 中央教育審議会 (2021).
- (5) 玉田周太郎, 原田恵雨, 佐藤奈々恵: “BERT を用いた TOEIC 問題自動生成システムにおける意味問題の生成”, 人工知能学会資料, 先進的学習科学と工学研究会, SIG-ALST-094-04 (2022).
- (6) 田中英武, 井垣宏, 嶋川一真, 福安直樹, 松本健一: “シナリオを活用した多様なプログラミング演習問題自動生成手法の提案”, 情報処理学会論文誌, Vol.66, No.4, pp.687-700, 情報処理学会 (2025).
- (7) 石田崇, 雲居玄道, 小林学, 平澤茂一: “生成 AI を用いた統計学の学習用練習問題自動生成の試み”, 日本教育工学会 2023 年秋季全国大会講演論文集, 16-2-S-002, pp.159-160, 日本教育工学会 (2023).
- (8) 松居辰則, 平嶋宗: “学習課題・問題系列のデザイン”, 人工知能学会誌, 25(2), pp. 259-267, 人工知能学会 (2010).
- (9) 平嶋宗, 東正造, 柏原昭博, 豊田順一: “補助問題の定式化”, 人工知能学会誌, 10(3), pp. 413-420, 人工知能学会 (1995).
- (10) 平嶋宗, 梅田多一, 志岐隆弘, 竹内章: “XML を用いた算数の文章問題の作成・共有環境: つるかめ算等の特定の計算手順で解決可能な文章問題を対象として”, 教育システム情報学会誌, 18(3・4), pp. 284-296, 教育システム情報学会 (2001).
- (11) 佐久間大, 金高右京, 後藤匠, 長谷川勝久, 佐藤隆博: “ISM 法と学習要素タグを用いた問題項目の構造化法の開発”, 日本教育工学会論文誌, 47(Suppl.), pp. 197-200, 日本教育工学会 (2024).
- (12) 佐藤隆博: “ISM 法による学習要素の階層的構造の決定”, 日本教育工学雑誌, 4(1), pp. 9-16, 日本教育工学会 (1979).