

修士論文

漢方医学的問診と共感力向上のための バーチャル患者モデルの開発

Development of a Virtual Patient Model for Kampo Medical
Interview and Enhancement of Empathy

富山県立大学大学院 工学研究科 電子・情報工学専攻

2355014 高田知樹

指導教員 António Oliveira Nzinga René 講師

提出年月: 令和7年(2025年)2月

目次

図一覧	ii
表一覧	iii
記号一覧	iv
第1章 はじめに	1
§ 1.1 本研究の背景	1
§ 1.2 本研究の目的	2
§ 1.3 本論文の概要	2
第2章 生成 AI の教育への活用	4
§ 2.1 教育に対する生成 AI の活用	4
§ 2.2 医療分野における生成 AI の現状と課題	8
§ 2.3 バーチャル患者の意義と生成 AI の役割	11
第3章 共感力向上のためのバーチャル患者	15
§ 3.1 対話システムの構築	15
§ 3.2 生成 AI モデルの学習手法	19
§ 3.3 問診と共感力	22
第4章 提案手法	27
§ 4.1 漢方医学的問診とバーチャル患者	27
§ 4.2 感情推定と表情の選択	30
§ 4.3 提案システムのアルゴリズム	35
第5章 数値実験並びに考察	40
§ 5.1 数値実験の概要	40
§ 5.2 実験結果と考察	41
第6章 おわりに	45
謝辞	46
参考文献	47

図一覧

2.1	自由研究おたすけ AI [19]	6
2.2	GDLS [20]	6
2.3	生成 AI の学習 [21]	7
2.4	情報の偏り [21]	7
2.5	恵寿総合病院と Ubie の事例 [23]	9
2.6	株式会社 HOKUTO の事例 [24]	9
2.7	Body Interact のシステムの流れ [29]	15
3.1	音声認識の種類 [33]	17
3.2	Whisper の流れ [34]	17
3.3	プロントエンジニアリングの役割 [41]	22
3.4	プロンプトの種類 [41]	22
4.1	トレーニングデータの内容	29
4.2	感情推定のプロンプト	32
4.3	感情の強度に対する表情	33
4.4	提案システムの流れ	37
4.5	実際のシステムの推移	38
5.1	出力結果	42

表一覧

2.1	医療教育に対する生成 AI を用いる課題と解決策	10
2.2	医療教育に対する生成 AI を用いる課題と解決策	11
2.3	模擬患者と標準模擬患者の違い [27]	13
3.1	Transformer と従来手法の違い	20
3.2	それぞれの尺度の違い	24
3.3	それぞれの尺度の違い	25
3.4	EES の項目の一部	26
4.1	問診シナリオの一部	28
4.2	モデルの比較	30
4.3	感情推定の種類	31
4.4	JSE の一部 [61]	35
4.5	ChatGPT の設定項目	36
4.6	VOICEVOX の設定項目	36
5.1	使用したシステムの環境	41
5.2	JES の結果	43
5.3	t 検定の結果	43

記号一覧

以下に本論文において用いられる用語と記号の対応表を示す.

用語	記号
事前学習の目的関数	$L_1(U)$
入力トークンのシーケンス	U
最後のトークン	u_n
トークンの確率	$P(u_i \mid u_{i-k}, \dots, u_{i-1}; \boldsymbol{\theta})$
次のトークン	u_i
コンテキストウィンドウのサイズ	k
モデルの学習可能なパラメータ	$\boldsymbol{\theta}$
対数確率の合計演算	$\sum_i$
トークン埋め込み	$\mathbf{h}_0$
入力トークンのシーケンス	U
トークン埋め込み行列	$\mathbf{W}_e$
位置埋め込み行列	$\mathbf{W}_p$
Transformer 各層の処理	transformer block()
第 l 層の出力	$\mathbf{h}_l$
Transformer モデルの層数	l
次のトークンの確率計算	$P(u)$
Transformer の最終層の出力	$\mathbf{h}_n$
トークン埋め込み行列の転置	$\mathbf{W}_e^T$
教師ありファインチューニングの確率計算	$P(y x_1, \dots, x_m)$
入力トークン列	$x_1, \dots, x_m$
正解ラベル	y
Transformer の最終出力	$\mathbf{h}_m$
タスク特化の線形出力層の重み行列	$\mathbf{W}_y$
教師あり学習の目的関数	$L_2(C)$
ラベル付きデータセット	C

はじめに

§ 1.1 本研究の背景

近年、世界的に補完代替医療の利用が増加しており、日本において特に信頼され、広く利用されているのは漢方医学である [1]. 漢方医学は西洋医学とは異なり、すべての患者に対して主訴の内容やその程度にかかわらず、伝統医学的な方法に基づいた問診、腹診、脈診、舌診などの特徴的な診察を行い、それらを総合的に判断して漢方医学的診断を行う. 正確な漢方医学的診断を行うためには、漢方医学的病態に基づいた問診を通じて患者の主観的かつ多岐にわたる訴えを詳細に聴取することが重要である.

さらに、この過程では、患者の表情を観察しつつ、患者の体験をその視点から理解し、共感を示すことが求められる [2]. 医師が患者に共感を示すことは、治療的な対人関係の構築に寄与し [3], 患者から正確で信頼性の高い情報を引き出すことを容易にする. このような関係性が構築されることで、より正確な診断が可能となり [4], 患者満足度の向上 [5] や健康アウトカムの改善が期待される [6].

医学生が問診を実施する機会は主に漢方臨床実習に依存しているが、COVID-19 パンデミックに伴う感染拡大防止策として、臨床実習に制限が課されていた. 現在、COVID-19 は感染症分類において5類に移行された. しかし、ウイルスは変化し続けており、変異株により毒性が高まる可能性は否定できないことが指摘されている [7]. そして、パンデミックが発生した場合には、臨床実習の実施が再び制限される可能性があるだろう.

このような状況から、医学生が患者に共感を示しつつ、必要な情報を効果的に引き出し、漢方医学的病態を理解するための問診教材の開発が急務であると考えられる.

近年、バーチャル技術の進歩に伴い、特にバーチャル患者を用いた医学教育が広く注目を集めている [8]. バーチャル患者と医学生に関する先行研究では、バーチャル患者とのコミュニケーションが実際の人間とのコミュニケーションに匹敵する感情的効果を有しており [9], バーチャル患者との相互作用を通じて医学生の共感が向上すること [10] [11] [12], さらに医学生の知識や診断精度の向上に寄与することが報告されている [13].

これらの知見から、漢方医学的問診におけるバーチャル患者の開発は、医学生の共感力を向上させたり、漢方医学的病態の理解を促進したりする可能性が示唆される. しかし、これまでのところ、バーチャル患者を用いた漢方医学教育に関する研究は行われていない.

§ 1.2 本研究の目的

1.1 節で述べたように、医学教育において、バーチャル患者の活用が注目されている。本研究では、バーチャル患者を開発し、教育効果の検証を目標とする。

バーチャル患者とは、頻繁に見られる臨床ケースから緊急事態に至るまで、幅広い実際の臨床シナリオをシミュレーションできる標準化されたコンピュータソフトウェアである。これを活用することで、実際の臨床現場では遭遇しにくい、または対応が難しい状況に対して、学生が安全に体験を積むことが可能となる [14]。

本研究で開発するバーチャル患者は、音声を通じた対話が可能であり、事前に設定したシナリオに基づいて回答を行う。また、感情推定を実施し、その推定結果に応じた表情変化をリアルタイムで反映させる機能も搭載する。これにより、学生はより現実的な問診体験が可能となる。さらに、本研究ではバーチャル患者の作成に生成 AI を活用する。

生成 AI とは、テキスト、画像、動画、音声など、多岐にわたるコンテンツを新たに生成できる人工知能 (AI) のことである。従来の AI が定められた処理の自動化を主眼としていたのに対し、生成 AI はデータから学習したパターンや関係性を活用して新たなコンテンツを生み出す点で革新的なものである。[15]

また、生成 AI は教育分野で活用されており、この技術を用いることで、効率的かつ柔軟にバーチャル患者の構築が可能となると考える。

漢方医学における問診訓練のためにバーチャル患者を開発することは、医学生にとって安全で実践的な学習環境を提供するだけでなく、漢方臨床実習の臨場感を維持しながら、問診スキルを繰り返し練習できる機会を提供することができるようになる。このような取り組みは、医学生の共感能力を高め、漢方医学的病態に対する理解を促進することが期待される。

また、文部科学省が策定した医学教育モデルコアカリキュラムに漢方医学が正式に組み込まれたことにより、従来の見習い制度に依存していた漢方医学教育の標準化が進んでいる [16]。しかし、指導者不足などの課題が依然として残っており、十分な標準化が達成されているとは言えない状況である。このため、バーチャル患者の開発は、漢方医学教育の標準化を支援する有効な手段となる可能性がある。

本研究の目的は、漢方医学的問診に特化したバーチャル患者を開発し、その教育効果を検証することである。具体的には、漢方治療を受けた患者の実体験を基にシナリオを作成し、そのシナリオを用いてバーチャル患者を設計を行う。さらに、このバーチャル患者を用いた学習を通じて、漢方医学的問診を適切に行えるのか、共感能力の向上の効果を検証を行う。研究の実用性を評価するために、学生を対象とした行動実験を実施する。この実験では、バーチャル患者に対して問診を行ってもらい、バーチャル患者の医学的問診の精度を計るとともに、その前後でアンケートを行い共感力を測定することで、本研究のバーチャル患者が共感力向上に寄与するかを検証する。

§ 1.3 本論文の概要

本論文は次のように構成される。

第 1 章 本研究の背景と目的について説明する。背景では、近年の漢方医学の重要性や、実

習教育の現状を説明している。目的では、漢方医学教育に対するバーチャル患者の開発の必要性を説明するとともに、本研究の意義を述べた。

第2章 生成 AI がどのように教育に用いられているのか述べるとともに、現在医療教育に用いられているバーチャル患者が使用された事例を述べ、どのように生成 AI を本研究に活用するのか説明している。

第3章 本研究のシステムを構築に使用される、対話型システムに必要な技術について述べる。また、共感の定義や共感力を上げる取り組みについて述べる。

第4章 音声認識技術と音声合成技術を用いて、対話型システムを作成する。また、生成 AI を用いて漢方医学的問診と共感力向上のためのバーチャル患者を開発する方法を述べる。

第5章 第4章で述べた手法で、バーチャル患者を作成し、実際に使用してもらい、本システムの有効性を示す。

第6章 本論文における前章までの内容をまとめつつ、本研究で実現できたことと今後の展望について述べる。

生成 AI の教育への活用

§ 2.1 教育に対する生成 AI の活用

AI 技術が急速に進化する中、教育の現場でもその影響が大きく現れている。生成 AI や自動化された学習プランニングの導入により、教育の質は飛躍的に向上し、効率的な学習が可能となった [17]。そこで教育に生成 AI を用いるメリットと実際の導入事例をいくつか述べる [18]。

教育に生成 AI を用いるメリット

- 生徒の学力や興味にマッチするコンテンツ提供

生成 AI はそれぞれの生徒に対して学力や興味を分析し、カスタマイズされた教育コンテンツを提供することができる。また、個別の学習計画の作成することにより、生徒個人に合わせたペースでの学習ができ、効率的に学ぶことができる。このように、生徒一人ひとりに合った教材を提供することで、興味の持続や学習効果の向上に寄与する。

- 生徒の学習中のリアルタイムでのアドバイス

生成 AI を用いることで、生徒の学習中にも適切なアドバイスやサポートを行うことができる。このように、リアルタイムで学習の進捗を評価し指導を行うことで、生徒の理解度の向上や学習の障壁を即座に解決することができ、学習効果が向上する。

- 生徒の学習意欲の向上

生成 AI が提供する、個人に最適化された学習を行うことで、生徒は自分の好奇心にあった内容で学習することができ、学習意欲が自然と高まる。このような体験は、学びに楽しさがうまれ生徒の主体的な学習を促す。

- 高度な学習機会の提供

生成 AI を用いることで、地理的な制約にかかわらず、高品質な学習の機会を得ることが可能となる。これにより、都市部だけでなく遠隔地に住む生徒にも同じような教育の機会を提供することができ、平等な学習を行うことができる。

- データ活用による教育の質の向上

大量のデータを分析することにより、生成 AI は継続的に教育の質を向上することができる。また、教材の効果を評価したり、教育のプログラムを改善することで、生徒により良い学習機会を提供することができる。

- **教師の業務負担軽減**

生成 AI を導入することで、出席の管理や、採点などのルーチン業務を自動化でき、教師の負担を軽減することができる。これにより、教師の授業の質の向上や、生徒個人に寄り添った指導の時間を増やすことができる。

- **教育の低コスト化**

生成 AI を活用することで、教育の運営コストが減り、品質の高い教材を多くの生徒に手ごろな価格で提供することができる。このようなコスト削減により、多くの生徒に高品質な教育を受ける機会を与えることができる。

生成 AI の教育への活用事例

- **つくば市立みどりの学園義務教育学校の事例**

つくば市立みどりの学園義務教育学校は、「生成 AI パイロット校」として AI を活用した授業を実施している。中学社会科の授業では、生徒たちが Bing チャットを使用し、地域課題に関する質問をして、調査活動を行った。この授業では、AI が提供する情報を参考にしつつ、生徒たちは教科書を併用して情報の正確性を確認しながらニュース原稿の制作に取り組んだ。この取り組みにより、生徒たちは情報リテラシーや批判的思考力を向上させるとともに、AI を効果的に活用するスキルを習得し、学力の向上にもつながる可能性が見られた。

- **長崎北高校の事例**

長崎北高校では、英語学習の一環として対話型 AI 「ChatGPT」を活用している。生徒たちは、英作文の添削や長文読解のサポートに AI を取り入れることで、自分の弱点を把握し、学力向上に役立てている。AI を活用することで、文法や表現方法など、自分では気づけなかった点を瞬時に指摘されるため、学習効率が向上している。この取り組みの特徴は、AI を活用した授業の中で、生徒たちが自らその活用法を実践・検証し、利用に関するガイドラインの作成に取り組んでいる点である。生徒たちは AI の利便性だけでなく、その課題についても理解を深め、効果的な活用方法を考えることで、問題解決能力を養っている。このような活動は、AI との適切な付き合い方を考える上で貴重な学びの機会を提供している。

- **愛媛大学教育学部附属中学校の事例**

愛媛大学教育学部附属中学校では、授業の「振り返り」を効率化するために、対話型 AI 「ChatGPT」を試験的に導入している。生徒たちはタブレット端末を使い、授業で学んだ内容や疑問点を入力すると、AI が即座にフィードバックを提供する。これにより、従来教師が時間をかけて行っていたコメント作成作業が効率化され、教師の業務負担が軽減される。さらに、この取り組みでは、AI が提供したフィードバック



図 2.1: 自由研究おたすけ AI [19]

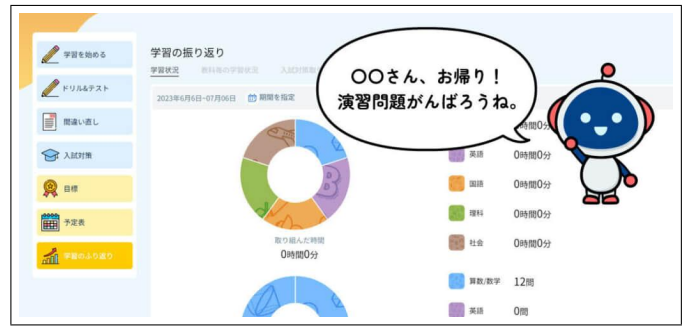


図 2.2: GDLS [20]

を教師が確認し、学習内容や生徒の理解度に応じて適切な修正を加えることで、教育の質を維持しながら業務効率化を実現している。このように、AIの利便性を活用しつつ、教師によるきめ細やかな対応を組み合わせたバランスの取れた活用方法は、生徒一人ひとりに対する丁寧な指導を保ちながら、教育現場の効率を向上させる。

● ベネッセの事例

ベネッセは、小学生とその保護者を対象に、「自由研究おたすけ AI」をリリースを行った(図 2.1)。このサービスは、生成 AI「ChatGPT」の技術を活用し、自由研究のテーマ選びをサポートするとともに、子供たちの疑問にアドバイスを提供する。利用者は、自由研究に使える時間や興味のあるジャンルを入力することで、AI「ラボリー」から具体的なテーマやアイデアの提案を受け取ることができる。この取り組みは、デジタルリテラシー教育の観点から保護者にも好評で、子供たちの学習を支援する新たな方法として注目を集めている。

● 学研の事例

学研ホールディングスは、オリジナル学習システム「GDLS」に ChatGPT を活用したベータ版を開始し、個別に最適化された学習アドバイスを提供している(図 2.2)。このシステムでは、生徒の学習履歴や理解度の変化を分析し、それに基づいた適切なアドバイスを提示することで、学習効果を最大化する。また、「GDLS」は、生徒が毎日ログインする習慣を促進し、学習意欲の向上を図ることができる。

このように、AI 技術の急速な進化により、教育の現場も大きな変革を遂げつつある。また、従来の一斉授業形式から脱却し、学習者一人ひとりに最適化された教育が実現しつつあり、「個別最適化学習」という新たなアプローチが注目されている。この方法では、AI が学習者の理解度や進捗状況をリアルタイムで分析し、それに基づいて個別化された学習プランを提供する。これにより、学習者は自分のペースで効率的に学ぶことができ、知識の定着度が向上する。

アメリカや日本などの教育先進国では、AI を活用した個別最適化学習の導入が進み、その有効性が実証されている。AI と教育の融合は、教育の現場に大きな変革をもたらし、未来の学習環境を構築する鍵として期待されている。

個別最適化学習の特徴は、AI 技術を駆使して学習者ごとに最適な学びを提供する点にある。このアプローチでは、AI が学習者の過去の学習履歴や現在のパフォーマンスを基に、

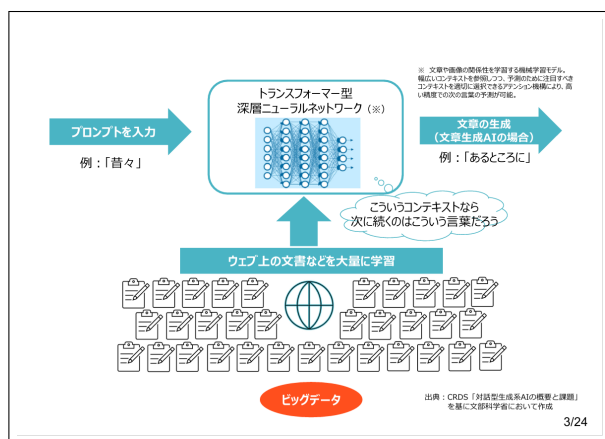


図 2.3: 生成 AI の学習 [21]

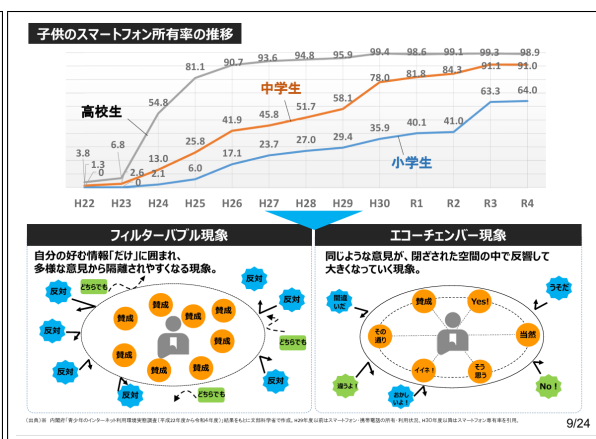


図 2.4: 情報の偏り [21]

次に学ぶべき内容や最適な学習ペースを提案する。これにより、学習者は理解不足や学習の遅れを解消しながら、効果的に知識を習得することが可能である。

さらに、個別最適化学習では、学習者の興味や関心に応じた教材の提供も行われる。例えば、特定の分野に興味を持つ学習者には、その分野に関連した深掘りコンテンツを提示することで、学習意欲を高める。このようなパーソナライズされた学習体験により、教育の質は飛躍的に向上する。

実際に、個別最適化学習を導入した教育機関では、学習成果や学習者の満足度が向上したとの報告がある。このアプローチは、従来の一斉授業では困難だった個別対応を可能にし、すべての学習者に最適な学習環境を提供する。さらに進化する AI 技術により、個別最適化学習は今後ますます普及し、教育の未来を大きく変える力を持つと考えられている [17]。

このように未来の教育環境は、AI 技術の進化によって大きく変わることが予想されている。しかし、生成 AI の活用は教育現場に多くのメリットをもたらすが、同時にいくつかの課題も存在する。文部科学省から使用するにあたってのガイドラインが出ており、生成 AI を用いることの課題を上げる [21]。

生成 AI 活用の課題

● 個人情報やプライバシーの懸念

生成 AI を利用する際、生徒や教師の個人情報が入力される可能性があり、適切に管理されない場合にはデータ漏洩やプライバシー侵害のリスクがある。教育現場では、個人情報を取り扱わない仕組みや、プライバシー保護のためのガイドライン整備が求められる。

● 誤情報やバイアスの問題

生成 AI は、図 2.3 のように学習するため、生成された情報には、不正確なデータや偏りが含まれることがある。このため、生徒が AI の出力をそのまま受け入れるのではなく、批判的に評価し、必要に応じて確認・修正するスキルが必要となる。そうしなければ、図 2.4 のような情報の偏りが埋めれてしまう。教師は生徒に対し、AI 利用時の情報リテラシーを育成する役割を果たさなければならない。

- **教育の公平性の確保**

AI技術の利用にはインターネット環境やデジタル端末が必要である。これらの設備が整わない地域では、AIを活用した教育の恩恵を受けにくくなる可能性がある。教育の公平性を保つためには、インフラ整備が不可欠である。

- **創造性への影響**

生成AIが簡単に情報や解答を提供することで、生徒の創造性や独自性が損なわれる可能性が指摘されている。教育現場では、AIを補助的なツールとして活用し、生徒の主体的な学びを促進する工夫が必要である。

生成AIは教育の質を飛躍的に向上させる可能性を秘めている。一方で、課題を克服し、安全かつ公平に利用するための対策が不可欠である。教育現場では、AIを補助的ツールとして活用し、創造性や主体性を育む指導方法を模索する必要がある。

生成AIを活用した教育の未来は、技術進化とともにさらなる可能性を広げていくと考えられる。その実現には、教師、生徒、保護者、そしてAI開発者が協力し、より良い学びの環境を構築することが求められる。

§ 2.2 医療分野における生成AIの現状と課題

高齢化の進行、慢性疾患の増加、医療コストの上昇など、医療業界はさまざまな課題に直面し、従来の医療体制には大きな負担がかかっている。このような状況の中で、持続可能な新しい解決策へのニーズが高まっており、生成AIの活用が医療分野においてますます重要になっている。そこで、医療分野で生成AIを用いるメリットと、実際の事例をいくつか述べる [22]。

医療分野で生成AIを活用するメリット

- **患者教育と情報提供の改善**

生成AIを活用することで、患者個々の病歴や症状に応じたカスタマイズされた健康情報やアドバイスを提供できる。これにより、患者が自身の状態をより深く理解し、適切な健康管理を行えるよう支援する。具体的には、生成AIが患者向けの説明文やアドバイスを生成し、わかりやすい形で情報を提供する。

- **医療文書の自動生成と整理**

診療内容や患者情報から自動的に診断書や治療計画などを生成することで、医師の文書作成にかかる負担を軽減する。生成AIが一部の文書を作成するだけでも、医師は患者ケアにより集中することが可能となる。

- **トレーニングと教育資料の生成**

疑似患者ケースやシナリオを生成することで、医療従事者のトレーニングを支援する。例えば、生成AIを用いて模擬患者を作成し、問診の練習や評価を行うことで、新人医師のスキル向上が図れる。

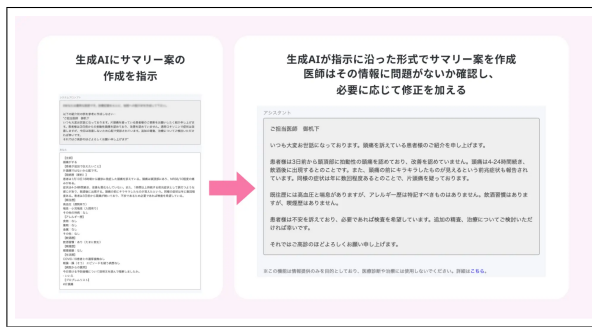


図 2.5: 恵寿総合病院と Ubie の事例 [23]



図 2.6: 株式会社 HOKUTO の事例 [24]

● 疾患の可視化とシミュレーション

生成 AI は、疾患の進行や治療の影響を視覚化する技術を提供することができる。画像生成 AI を活用することで、疾患の状態を模倣した画像を作成し、患者や医療従事者が病態をより深く理解できるようサポートする。

● カスタマイズされた治療計画の提案

生成 AI とデータ分析を組み合わせることで、患者ごとに最適な治療計画を提案することが可能となる。これにより、医療の質が向上し、患者のニーズに応じた高度なケアが提供される。

● 手術前に患者の 3D モデルを使ってシミュレーション

患者の CT や MRI データから生成した 3D モデルを用いて、外科医が手術のシミュレーションを行うことで、精度と安全性を向上させることができる。また、この技術は医学教育や患者への説明にも役立つ。

● 画像生成 AI による診断画像の画質改善

GAN（敵対的生成ネットワーク）技術を用いて、低線量 CT 画像を高画質化することで、放射線被ばく量を低減しながら、鮮明で詳細な画像を提供する。これにより、微小な病変の発見や正確な診断が可能となる。

医療分野における生成 AI の活用事例

● 恵寿総合病院と Ubie による文書作成負担の軽減

恵寿総合病院と Ubie 株式会社は、図 2.5 のように、生成 AI を活用して診断サポートや診療記録の自動生成に取り組んだ。この取り組みは、医師の退院時サマリー作成時間を最大で 1/3 に短縮し、患者とのコミュニケーションに割ける時間を増やすことに成功した。また、業務効率の向上だけでなく、医療サービスの質の向上も確認されている。

● Hippocratic AI による医療プロセスの包括的サポート

Hippocratic AI は、診断支援、手術計画、症状チェック、薬の服用確認など、幅広い医療プロセスを支援する AI ソリューションを提供している。この技術は、患者の生活習慣改善や慢性疾患管理、退院後のケアに焦点を当てており、作業負担の軽減と効率的な患者管理を実現すると同時に、患者ケアの質を高める手段として機能している。

表 2.1: 医療教育に対する生成 AI を用いる課題と解決策

分野	効果
模擬患者シナリオや疑似ケースを生成	実践的なトレーニング環境を提供
教育資料やトレーニング教材の自動生成	学習時間を短縮しつつ内容の質を高める
多言語対応の教材を容易に生成	国際的な医療教育に大きな価値を発揮
医療教育における評価プロセスの効率化	個々の弱点や改善点を特定

● Insilico Medicine による医薬品開発の効率化

Insilico Medicine は、生成 AI を用いて新しい抗線維化薬「ISM001-055」を開発した。このプロセスでは、膨大なデータから治療ターゲットを特定し、候補となる化学構造を設計。通常よりも短期間かつ低コストで薬剤開発を進め、早期臨床試験に移行することが可能となった。

● BlueMeme と九州大学による量子 AI 活用

BlueMeme と九州大学は、量子 AI を使用してバイオメディカル分野に特化した言語モデル（LLM）の共同研究を行っている。この研究では、医療研究や疾患原因探索、薬剤設計に役立つ知見を提供するツールの開発を進めており、量子 AI による運用コストの削減も目指している。

● GPT-4 を活用した医師支援（株式会社 HOKUTO）

株式会社 HOKUTO は、図 2.6 のような臨床支援アプリに OpenAI GPT-4 を組み込んだ新機能を導入した。この機能では、患者への説明案の生成やキーワードに基づく最新研究論文の抽出が可能となり、医師の情報収集負担を軽減する。多言語対応の患者説明案生成機能により、患者理解と信頼関係の構築が促進されている。

生成 AI の医療分野での活用は、診断支援や個別化された治療計画の作成、遠隔医療や予防医療の促進など、多岐にわたる。これらの技術は、患者ケアの質を向上させるだけでなく、医療や開発コストの削減、医療アクセスの拡大にも寄与している。このように、生成 AI の医療分野への応用は非常に重要であることがわかる。また、医療教育の分野でも生成 AI の活用方法が表 2.1 のように考えられている。

まず、生成 AI の医療教育分野への応用は、医療従事者のスキル向上や学習効率の向上に寄与する可能性を秘めている。具体的には、模擬患者シナリオや疑似ケースを生成することで、実践的なトレーニング環境を提供できる。例えば、新人医師が問診や診断のスキルを磨く際、生成 AI を用いて作成された模擬患者が多様な症状や病歴を持つケースを再現することで、現実に近い学習体験を提供できる。これにより、実際の患者を相手にする前に必要なスキルを磨く機会を得られるだけでなく、教育者にとっても効率的な指導が可能となる。

さらに、生成 AI は、教育資料やトレーニング教材の自動生成にも役立つ。医療従事者向けの最新の研究成果を要約した資料や、特定の疾患に関する診断ガイドラインなどを迅速に生成することで、学習時間を短縮しつつ内容の質を高めることができる。また、多言語

表 2.2: 医療教育に対する生成 AI を用いる課題と解決策

課題	具体例	対策
コンテンツの信頼性と品質保証	誤情報や偏見が含まれる可能性	データセットの厳選、第三者によるコンテンツの検証プロセス
プライバシー保護	患者データの利用によるプライバシー侵害リスク	データの匿名化、透明性の確保
学術的不正や依存のリスク	学生がAIに過剰依存し、スキルを学習できなくなる	AI利用の適切な方法を教育
法的な不整備	AIが生成したコンテンツの責任の所在が不明確	医療AIに関する規制の整備と国際ガイドラインの作成

対応の教材を容易に生成することも可能で、国際的な医療教育の場においても大きな価値を発揮する。

加えて、医療教育における評価プロセスの効率化も期待される。生成 AI を活用すれば、医学生のパフォーマンスを分析し、個別のフィードバックを自動的に提供することが可能となる。これにより、従来の画一的な評価方法では見逃されがちだった個々の弱点や改善点を特定し、より個別化された教育が実現できる。また、生成 AI によるリアルタイムのフィードバックは、学習者がその場で改善策を実践し、効果を確認するサイクルを促進するため、学習の定着率を高めることができる。

一方で、生成 AI を医療教育に導入する際には、表 2.2 のように、いくつかの課題にも対処する必要がある。特に、生成されたコンテンツの信頼性と品質をどう保証するかが重要である。医療教育の場では、正確で現実的な情報が不可欠であり、AI が偏見や誤った情報を生成する可能性に備える必要がある。このため、生成 AI モデルのトレーニングに使用されるデータセットの厳選や、コンテンツの検証プロセスの確立が求められる。

また、倫理的な配慮も重要である。生成 AI が学術的不正や過剰な依存を助長しないよう、学生や教育者に対して AI の適切な利用方法を教育することが不可欠である。さらに、個人情報保護の観点から、患者データを AI モデルに使用する際には、厳格なプライバシー保護対策が求められる。これには、匿名化されたデータの使用や、データ処理の透明性の確保が含まれている。

将来的には、生成 AI が医療教育に与える影響を定量的に評価し、その効果を最大化するためのベストプラクティスを確立することが求められる。継続的な研究と学際的な協力を通じて、医療教育の質を高めるだけでなく、医療現場で即戦力となる人材の育成にも寄与することが期待される。生成 AI の技術が進化する中で、教育者や研究者、実務家が協力し、その潜在能力を最大限に活用する取り組みがますます重要になる [25]。

§ 2.3 バーチャル患者の意義と生成 AI の役割

1.1 節で述べたように、医学生が問診を実施する機会は主に漢方臨床実習に依存している。ここで、臨床実習とはどのようなものなのか述べる。

臨床実習とは、病院で行われる実習のことで、その期間は大学や学科によって異なる。一般的には、1 年次に 1 週間程度、4 年次に 4 週間程度の期間で実施される。一方で、医学部

や歯学部では、合計で約2年間にわたる臨床実習が行われる。医学部の臨床実習には、早期体験実習や解剖実習、見学型実習、そして診療参加型実習といった多様な形式が含まれており、以下でいくつか説明する [26]。

臨床実習の種類

● 早期体験実習

早期体験実習は、見学型臨床実習や参加型臨床実習に先立って、1～2年次に行われる臨床医学の導入的な実習の1つである。その目的として、コミュニケーションスキルの向上、医学への学習意欲の向上、患者の心情や視点への理解がある。実習では、外来患者のエスコートや病院内の各部署での業務体験が行われる。これにより、医療現場で貢献する達成感を得られ、学ぶモチベーションが高まる。また、病院全体の仕組みや業務の流れを理解する機会にもなる。

● 解剖実習

解剖実習は、医学部の学生が4～5人のグループで実際のご遺体を用いて行う実習である。臓器や筋肉、神経の位置や機能を実際に観察し、座学で学んだ知識を実際の人体で確認することで理解を深める。この実習は人体の構造や機能を理解するだけでなく、解剖技術の向上を目指すものである。さらに、ご遺体に向き合い、人間の身体に手を加えることの重大さや倫理的意義を考える重要な機会でもある。

● 見学型臨床実習

見学型臨床実習は、高学年の医学部生が医療機関で行う実習で、主に4～6年次に実施される。指導医のもとで外来や手術を見学し、臨床の現場を学ぶことを目的としている。病棟実習では、患者を担当することもあり、診察や治療を通じて疾患をさまざまな観点から学ぶことができる。

● 参加型臨床実習

参加型臨床実習は、医学生が特定の診療科を選択し、一定期間その診療科の医療チームに加わり、実践的な臨床能力を磨く実習である。医学生は「スチューデント・ドクター」として指導医とともに診療に参加する。最終的な診断や治療方針の決定は行わないが、患者への問診や診察を自分で行い、診断や治療について考える経験を得られる。この実習は、医師免許取得後の初期臨床研修とのギャップを埋める目的でも実施されている。

このように、様々な臨床実習が行われおり、医学教育において臨床実習が重要であることがわかる。しかし、近年、医療を取り巻く社会情勢の変化が医学教育に大きな影響を与えている。医療の高度化・複雑化や入院期間の短縮など、医療の形態が変化する中で、医療安全や患者倫理への配慮が求められ、実際の患者を用いて練習する機会が減少している。これに対応するため、2000年代以降、シミュレーション教育が導入され、患者との直接的な臨床実習の前段階で、繰り返し練習ができる安全な教育環境が提供されるようになった。特に、リアルな人間とのシミュレーションを用いた教育として、模擬患者（Simulated Patients /

表 2.3: 模擬患者と標準模擬患者の違い [27]

	模擬患者(simulated patient : SP)	標準模擬患者(standardized patient : SP)
目的	コミュニケーションの練習	試験への参加
役割	演技とフィードバック	演技と評価
活動場面	授業、実習、研修会など	OSCE (objective structured clinical examination ; 客観的臨床能力試験)
シナリオ	比較的簡単なシナリオ	一問一答の詳細なリスト
演技	気持ちの動きに従って自然な会話をする比較的自由度のある演技	SP 間で演技を標準化し、会話のルールに従った演技
評価	なし	標準化された基準に沿った評価

Standardized Patients, SP) を活用したコミュニケーション教育の重要性が高まっている。2022 年度に改訂された医学教育モデル・コア・カリキュラムでは、コミュニケーション能力の育成が強調されており、「患者およびその関係者と良好な関係性を築き、相手の状況を考慮しながら意思決定を支援し、安全で質の高い医療を実践する」ことが求められている。このカリキュラムには、模擬患者、身体部分模型、そして VR (Virtual Reality) などの先進的な教育手法の活用が含まれている。また、2019 年度の検討報告書では、近年の人間関係の希薄化や生活体験の不足が進行している状況を受けて、コミュニケーション能力向上のための教育をさらに強化する必要性が指摘されている。現在、模擬患者 (SP) は医療系大学における医療面接教育の中で、標準的な人的教育資源として広く活用されている [27]。

Simulated patients (SP) とは、実際の患者を忠実に再現するために十分な訓練を受け、熟練した医師でも見分けがつかないほど高い精度で患者役を演じることができる個人を指す。SP は単に病歴を伝えるだけでなく、身体の動きや身体所見、さらに感情や性格までも含めた患者像を再現する役割を担う。SP には、模擬患者 (simulated patients) と標準模擬患者 (standardized patients) の 2 種類があり、表 2.3 のような違いがある。

このように、臨床実習に欠かすことのできないものであるが、sp にはいくつかの課題がある [27]。

SP の課題

● 配慮ある環境を設定し、身体診察に協力可能な SP をリクルートする必要

身体診察に協力可能な SP を確保するためには、安全で配慮の行き届いた環境を整備することが求められる。臨床実習後の OSCE では、1 人の SP が診察を受ける形が理想的だが、現在はシミュレーターの活用や、診察時に女性 SP と男性 SP を交代する方法が用いられる場合もある。しかし、よりリアルな状況を再現するには、1 人の SP が診察を受ける形が望ましい。全国調査では、SP が身体診察への意義を感じており、年齢や性別、診察部位を考慮すれば協力が得られる可能性が示されてる。また、2019 年度から OSCE で女性の下着着用が認められ、女性 SP の参加が増加していることも明らかになった。身体診察に対する羞恥心や不安を軽減し、安心して診察を受けられる環境を構築することが重要な課題である。

- **SP 養成が短期間では完了せず、根気強く継続的な取り組みが求められる**

SP 養成が短期間では完了せず、根気強く継続的な取り組みが求められる。そのため、SP 養成者は SP との練習時間を十分に確保する必要があります。この準備練習は、演技やフィードバックの質に直結する重要なプロセスです。しかし、OSCE における標準 SP 養成業務の負担が増加している現状では、非常勤ポストを設けて専任の SP 養成者を確保することが大きな課題となっている。

- **設定標準 SP 制度の導入に伴い、SP 養成の質をさらに向上させる必要**

令和5年度から OSCE が公的試験化された背景には、医師法の改正により臨床実習で医業が可能になったことがある。この制度では、公的 OSCE に対応できる標準化された SP の演技が求められ、規定された演技と評価をクリアすることが必須である。このため、質の高い演技指導ができる SP 養成者の育成が喫緊の課題となっている。

このように SP には、様々な課題があるわけだが、そんな中、1.1 節でも述べたように、バーチャル技術の進歩に伴い、模擬患者の代わりにバーチャル患者を用いた医学教育が広く注目を集めている [8]。例えば、Body Interact¹というアプリケーションがある。Body Interact は、医療教育のための仮想患者シミュレーターであり、安全な環境でリアルな臨床シナリオを体験できるツールの一例である。学習者が行う診断や治療に応じて患者の状態がリアルタイムで変化し、現実的な医療状況が再現される。このシステムでは、感染症や心肺停止、外傷など多様なシナリオを通じて、学習者が臨床推論スキルを向上させることが可能である。

さらに、Body Interact はリアルタイムフィードバック機能を備えており、学習者が行った診断や治療が患者に与える影響を即時に確認できる。例えば、不適切な治療が患者の状態を悪化させる場合、その結果を直感的に理解することで、誤りを修正しながら学習を進めることができる。このような即時フィードバックは、学習効果を高め、スキルの定着を促進する。

また、個別学習とチーム学習の両方に対応している。個別学習では、自分のペースで取り組むことが可能であり、自己評価に集中できる。一方、チーム学習では、複数の学習者が協力してシナリオを進めることで、コミュニケーションやチーム医療のスキルを向上させる機会を提供している (図 2.7)。

さらに、多言語対応や複数デバイスでの利用が可能である点も大きな利点である。これにより、言語の壁を超え世界中の学習者が利用できる教育ツールとして広く導入されている。

このように、Body Interact は、従来の SP の課題を補完しつつ、学習者に多様な教育機会を提供する新しいツールとして、その価値を高めている。特に、AI 技術の統合により、学習者が効果的に臨床スキルを習得し、実際の医療現場での応用力を高めることが期待されている。今後、このようなツールは医学教育において重要な役割を果たすとかんがえられる。

また、バーチャル患者に生成 AI を用いる研究も近年行われている [28]。医療教育におけるシミュレーション技術の進歩により、学生が安全な環境で実践的なスキルを学ぶ機会が広がっている。本研究では、OpenAI が開発した生成 AI モデル「ChatGPT」を仮想患者と

¹<https://bodyinteract.com/>

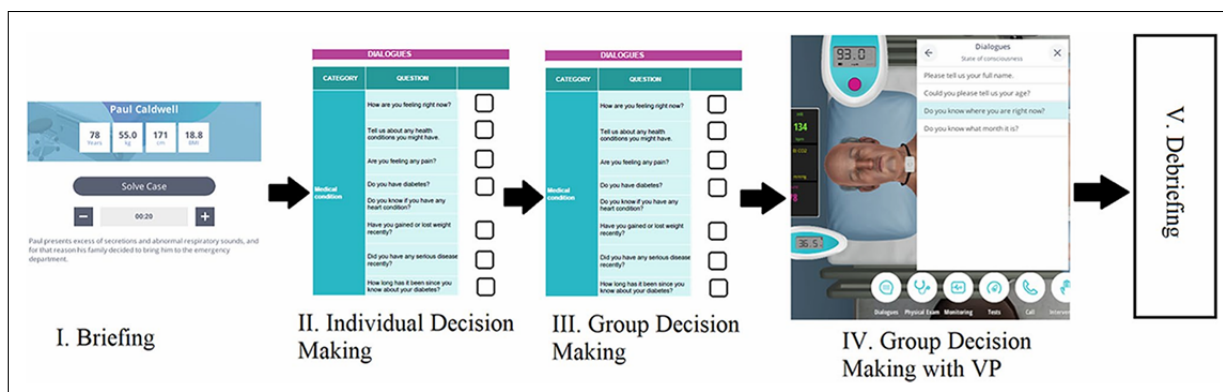


図 2.7: Body Interact のシステムの流れ [29]

して活用し、看護教育におけるトレーニング効果を評価した。この研究は、モロッコの高等健康科学研究所に所属する 12 名の看護学生を対象に行われ、呼吸困難の患者シナリオを用いた対話型シミュレーションを実施したものである。

研究の評価項目として、ChatGPT の受容性、アクセシビリティ、エンゲージメント、そして学生の仮想患者とのインタラクションスキルが挙げられた。結果として、ChatGPT は学生から高く評価され、ツールの使いやすさや没入度が特に優れていることが分かった。また、学生のインタラクションスキルは、回答の明瞭さや関連性、有用性といった点で向上が見られた。特に、回答の明瞭さと理解度には強い相関が認められ、シミュレーションの効果が確認された。

ChatGPT を活用することで、学生は対話を通じて実践的なコミュニケーションスキルを習得できる。従来のシミュレーションと比較して、ChatGPT はコスト効率が高く、再現性や適応性に優れる点で優位性を持つ。しかし、一方で、時折不適切な回答を生成する可能性があるため、教育者による監視や補完が必要であることも指摘されている。

この研究は、看護教育における ChatGPT の可能性を示しただけでなく、仮想患者シミュレーションが学生の学習体験をどのように変革し得るかを明らかにした。ChatGPT を活用したシミュレーションは、学生が患者とのコミュニケーションスキルを磨くと同時に、実際の医療現場で求められる準備を効果的に進める手段となり得る。このような AI 技術を教育に統合することで、従来のトレーニング手法を補完し、新しい学習の形を提案することが可能である。

結論として、ChatGPT を仮想患者として用いたシミュレーションは、看護学生の臨床準備を強化し、効果的な学習環境を提供するツールとしての可能性を示した。これにより、学生は実際の患者との対話に向けた準備を進めることができ、看護教育全体の質を向上させることが期待されている。

共感力向上のためのバーチャル患者

§ 3.1 対話システムの構築

近年の人工知能ブームの影響により、Siri に話しかけたり、店頭に設置された Pepper のようなロボットと会話したり、さらには Line 上でりんなどの対話を楽しむ人々が増えてきた。また、商業的な視点では、コンタクトセンターなどの問い合わせ対応業務の負担を軽減する目的で、チャットボットを導入する企業が増加している。これにより、従来は人間のオペレーターが行っていた対応の一部を自動化し、効率化を図る動きが主流となりつつある [30]。そこで対話システムとは、どのようなものなのか述べる。

対話システムは、人間とコンピュータが自然な言葉で円滑に会話するために研究が進められており、タスク指向型と非タスク指向型の2つに分類される [31]。

タスク指向型システム

タスク指向型システムは、ユーザーの質問や要求に応じて必要な情報を提供することを目的としている。例えば、Apple の「Siri」や NTT ドコモの「しゃべってコンシェル」などがあり、音声を使ってスケジュール管理や Web 検索などを行う。

非タスク指向型システム

非タスク指向型システムは、特定の目的がない自由な会話を通じて、ユーザーに楽しさや自然な対話を提供することを目的としている。代表的なシステムとして、初期の「ELIZA」やその発展形である「A.L.I.C.E.」が挙げられる。

この応答パターンの代表的な作成手法には、ルールベース、抽出ベース、生成ベースの3種類がある [32]。

ルールベース

ルールベースは、あらかじめ人が応答ルールを設定する方法である。たとえば、ユーザーが「こんにちは」と発話すればシステムが「こんにちは」と返し、「眠い」と言えば「おやすみ」と返すように、想定されるユーザーの発話に対して手動で応答を定義する。ルールベースは手動で設計するため、応答内容の管理が容易で、不適切な発話が少ないという利点がある。また、時事ネタなどユーザーの興味を引く発話を追加しやすい点も強みである。このため、Siri や Alexa といった現在の商用対話システムの多くは、ルールベースが採用されている。しかし、人手でルールを作成するため、幅広い話題に対応することは難しいという欠点がある。

抽出ベース

抽出ベースは、大量のデータから適切な応答例を検索して利用する方法である。具体的には、ユーザーの入力に類似した単語や文を見つけて返す「類似文検索」と、対話ログやQA データなど応答ペアを活用する「ペア利用」の2つのアプローチがある。構築コストが低く、幅広い話題に対応可能という利点があるが、類似文検索では単なるオウム返しになることが多く、ペア利用では文脈がずれて関連性の低い応答が出る可能性があるという欠点がある。

生成ベース

生成ベースは、用例を基にしながらも応答の品質を向上させる手法である。大量のテキストデータを事前に学習し、文法や係り受けの関係を抽出することで、ユーザーの発話内容に関連する話題やキーワードを選び、自然で高品質な応答を生成する。この手法では、抽出ベースの欠点であった「無関係な発話」や「オウム返し」の発生を抑えることができる。

また、対話システムには、チャットボット (Chatbot) とボイスボット (Voicebot) の2種類がある。チャットボットは、テキストベースのインターフェースを介してユーザーと対話を行うのに対し、ボイスボットは音声を用いて対話を行うシステムである。本研究では、より自然で直感的な対話を実現するために、ボイスボット形式を採用する。ここでは、ボイスボットを構築するために必要となる主要な技術について詳しく説明する。

ボイスボットでは、チャットボットとは異なり、主に音声認識技術とテキスト音声合成技術の2つの技術が使用されている。

音声認識技術

音声認識技術は、ユーザーの話した音声をテキストデータに変換する技術である。これにより、システムは音声入力を理解し、適切に応答することが可能となる。音声認識の一般的な処理の流れは以下の通りである [33]。

1. 音声入力

マイクやその他の入力デバイスを用いて、認識対象の音声を録音または入力をおこなう。

2. 音響分析 (音声のデジタル化)

入力された音声信号から特徴量を抽出し、デジタルデータに変換する。この過程では、時間領域や周波数領域の情報を活用する。

3. 音波から音素の特定 (音響モデル)

音響モデルを使用して、音声データから音素 (言語の最小単位) を特定する。このモデルは、機械学習アルゴリズムを利用して音波の特徴を音素に対応付を行う。

4. 音素の並びから単語への変換

特定された音素の並びを解析し、文脈情報や言語モデルを用いて適切な単語や文に変換を行う。

5. テキスト出力

最終的に、認識結果をテキスト形式で出力する。この出力は、自然言語処理や翻訳システムなどに利用されることがある。

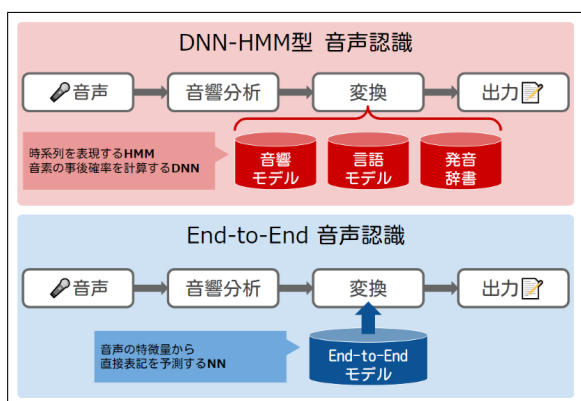


図 3.1: 音声認識の種類 [33]

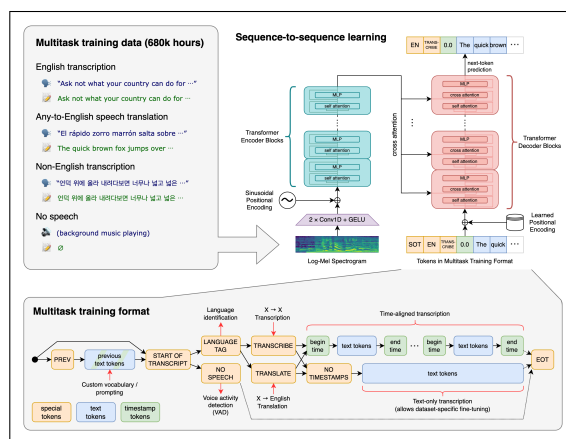


図 3.2: Whisper の流れ [34]

音声認識技術には大きく分けて図 3.1 のような 2 つのアプローチがある。それは、DNN-HMM 型と End-to-End 型である [33]。

DNN-HMM 型

音響モデル、言語モデル、発音辞書の 3 つのモジュールを組み合わせて音声テキストに変換する従来の手法である。音響モデルには、DNN（ディープニューラルネットワーク）と HMM（隠れマルコフモデル）を組み合わせたハイブリッド型が用いられている。この方法では、各モジュールの組み合わせやパラメータ調整が必要であり、システム全体が複雑になる一方、特定のタスクや専門用語への適応が比較的容易であるという利点がある。

End-to-End 型

単一のニューラルネットワークで音声から直接テキストを生成する新しい手法である。このアプローチでは、発音辞書を必要とせず、音声とその書き起こしデータのみで学習が可能で、システムの構築がシンプルであるというメリットがある。しかし、特定の単語や専門用語への適応が難しく、大量の学習データと高い計算リソースが必要とされる課題も指摘されている。

本研究では、音声認識に End-to-End のものを用いる。そこで End-to-End 音声認識の 1 つである、OpenAI が開発した Whisper について説明する。Whisper は、End-to-End 型の高度な音声認識システムであり、多様な言語やノイズ環境下でも高い精度で音声をテキスト化することが可能という特徴がある。次に、Whisper のシステムの流れについて説明する (図 3.2) [35]。

1. 音声データの入力

Whisper は、音声ファイル (例: WAV, MP3, FLAC) を入力として受け取り、すべての音声データは、16kHz のサンプリングレートに統一され、処理が開始される。

2. 特徴量抽出

入力された音声信号は、メルスペクトログラム (Mel Spectrogram) に変換される。メルスペクトログラムは、音声の時間-周波数情報を表す形式であり、人間の聴覚感度に

基づいて設計されている。入力音声は固定サイズ（30 秒）のメルスペクトログラムパッチとして処理される。

3. エンコーダ処理

エンコーダは、Transformer アーキテクチャをベースにしており、メルスペクトログラムを高次元の埋め込みに変換する。この処理により、音声信号の時間的および周波数的な特徴が抽出され、内容を理解するための基盤が構築される。

4. 言語モデリング（デコーダ処理）

デコーダは、音声信号からテキストを生成する役割を担う。言語モデルとしての Transformer デコーダが、音声特徴量と事前学習された言語知識を組み合わせ、適切なトークン列（文字列）を生成する。

5. 自動言語検出

Whisper は、入力音声の言語を自動的に検出を行う。この多言語対応モデルにより、100 以上の言語を認識可能である。言語検出後、その言語に最適化された処理を行う。

6. テキスト生成

デコーダは、音声特徴量に基づいて逐次的にトークンを生成し、最終的なテキストを出力する。この出力形式には、タイムスタンプ付きテキストやフォーマット済みのテキストが含まれる。

テキスト音声合成技術

テキスト音声合成技術は、基本的な音声合成技術の一つであり、生成されたテキストの応答を自然な音声に変換する技術である。これにより、ボイスボットはユーザーに音声で情報を提供したり、返答したりすることができます。この技術の利点は、多様な単語やフレーズを音声として生成できる点にある。テキスト音声合成は、大きく分けて、「規則ベースの合成方式」と「コーパスベースの合成方式」の2つに分類される [36]。

規則合成方式

あらかじめ設定された規則に基づいて音声波形を合成する方式である。新しい単語にも対応できるが、音声が機械的になりやすく、リアルな人間の声を再現するのは難しい。そのため、ウェブサイトやメールの読み上げなど、アクセシビリティツールとして限定的に使用される。

コーパスベース合成方式

事前に収集した会話、スピーチ、インタビューなどの音声データを基に、入力テキストに対応する音声波形を用いて音声を合成する方式である。この方式はさらに以下の2つに分類される。

1. 波形接続型音声合成方式

大量の音声データを収録し、音素単位に分割して結合することで音声を合成する方式である。高音質な音声を生成できるが、自然な音声を作成するには専門的な知識とスキルが必要であり、主にプロや上級者向けである。

2. 統計モデル型音声合成方式

音声データと機械学習を用いて音声モデルを構築し、そのモデルに基づいてテキ

ストの解析と音声予測を行う方式である。音声データが少なくても安定した音声合成が可能である。最近では、深層学習技術を活用した WaveNet や Tacotron などが、この方式に該当する。

§ 3.2 生成 AI モデルの学習手法

本研究では、生成 AI を用いて応答の生成を行う。生成 AI は自然言語処理を基盤としているが、これまでの自然言語処理では主に RNN や LSTM などの手法が用いられてきた。しかし、近年では Transformer が主流となり、広く活用されている。そこで、Transformer が従来の RNN や LSTM とどのように異なるのかについて述べる（表 3.1 参照）。

RNN

RNN (Recurrent Neural Network) は、時間的な依存性を持つシーケンスデータや時系列データの処理に適したニューラルネットワークの一種である。このネットワークの特徴は、各ステップの出力が次のステップの入力として再利用される再帰的な構造を持つことである。これにより、最新の情報だけでなく、過去の情報も統合して処理することが可能になる。つまり、RNN は過去の状態を考慮して次の状態を予測する能力を持っている。RNN の構造では、各ステップで入力データに対する重み行列と、前のステップの出力データに対する重み行列が使用される。この仕組みにより、RNN は時系列データの特徴やパターンを学習し、次のステップで適切な出力を生成することができる [37]。

LSTM

LSTM (Long Short-Term Memory) は、RNN の一種で、特に長期的な依存関係を持つデータを効果的に学習できる点で注目されている。従来の RNN が抱えていた「勾配消失問題」を解決するために開発された LSTM は、情報の流れを調整する「ゲート機構」を持っている。この仕組みにより、重要な情報を長期間保持し、不要な情報を効率的に「忘れる」ことが可能となっている。この特性により、LSTM は自然言語処理や時系列データの予測など、複雑な依存関係が求められる分野で優れた性能を発揮している [37]。

Transformer

Transformer は、Google が 2017 年に発表した自然言語処理のためのモデルである。従来の RNN を使わず、自己注意機構 (Self-Attention) のみを用いて、長文でも効率よく情報を処理する。特に、エンコーダとデコーダの 2 つの構成から成り、マルチヘッドアテンションや位置エンコーディングが特徴である。これにより、翻訳や文章生成などで高精度かつ高速な処理が可能になった [38]。

こうした技術的な進化に伴い、生成 AI を特定のタスクに適応させるための学習手法が注目されている。生成 AI にはさまざまな学習手法があり、その中でも代表的なものの一つがファインチューニングである。ファインチューニングとは、大規模なデータセットで事前学習された機械学習モデルを基盤とし、特定のタスクや分野に適応させるために調整を行うプロセスです。この手法には以下のような特徴がある [39]。

表 3.1: Transformer と従来手法の違い

特徴	Transformer	RNN/LSTM
アーキテクチャ	Self-Attentionを中心とした並列処理が可能な設計	リカレント構造に基づき、データを逐次的に処理する設計
メモリコスト	高い	低い
性能	大規模データで非常に高い性能を発揮	中規模データでは良好だが、大規模データには限界がある
長距離依存の処理	長距離依存を効果的に捉えられる	長距離依存を捉えるのが苦手

- **事前学習モデルの活用:** 大規模なデータセットで事前学習されたモデルを基盤として使用
- **特定タスクへの適応:** 対象とするタスクに関連した小規模なデータセットを用いてモデルを適応
- **パラメータの微調整:** モデル全体または一部のパラメータを再訓練して調整
- **転移学習の一環:** 一般的な知識を特定の問題領域に応用する手法

ファインチューニングにはさまざまな手法があるが、本稿では、事前学習（Pre-Training）から教師ありファインチューニング（Supervised Fine-Tuning, SFT）に至るプロセスを取り上げ、本研究のシステムに採用されている生成 AI にも活用されている SFT について説明する。

事前学習

まず事前学習として大規模なテキストデータセットを使用し、モデルが次の単語を予測できるように学習を行う。モデルは、入力として与えられたトークンの列を使用し、次のトークンが現れる確率を最大化するように学習を行う。素の学習には、以下の目的関数を用いる。

$$L_1(U) = \sum_i \log P(u_i | u_{i-k}, \dots, u_{i-1}; \theta) \quad (3.1)$$

ここで、 $U = \{u_1, u_2, \dots, u_n\}$ が入力トークンのシーケンスであり、 $P(u_i | u_{i-k}, \dots, u_{i-1}; \theta)$ がモデルが過去の k 個のトークンを基に次のトークン u_i を予測する確率、 θ がモデルの学習可能なパラメータ、 k がコンテキストウィンドウのサイズである。この目的関数は、各トークンの出現確率の対数を合計することで、モデルが正しい予測を行うように学習を進める。

入力トークン $U = \{u_1, u_2, \dots, u_n\}$ を、トークン埋め込み行列 $\mathbf{W}_e$ を用いて数値ベクトルに変換を行う。また、トークンの位置情報を考慮するために、位置埋め込み行列 $\mathbf{W}_p$ を加える。この処理は以下の式で表される。

$$\mathbf{h}^0 = U\mathbf{W}_e + \mathbf{W}_p \quad (3.2)$$

ここで、 $\mathbf{h}^0$ はモデルの初期状態である。初期状態 $\mathbf{h}^0$ を Transformer ブロックに入力し、self-attention を用いてトークン間の関係を学習を行う。この処理は以下のように表される。

$$\mathbf{h}^l = \text{transformer_block}(\mathbf{h}^{l-1}), \quad l \in [1, n] \quad (3.3)$$

ここで、 l は Transformer モデルの層数を表し、各層が前の層の出力を利用して計算を行う。

Transformer の最終層 $\mathbf{h}^n$ の出力を使い、次のトークン u_i が現れる確率を計算する。トークン埋め込み行列の転置 $\mathbf{W}_e^T$ を適用し、ソフトマックス関数で確率分布を出力する。

$$P(u) = \text{softmax}(\mathbf{h}^n \mathbf{W}_e^T) \quad (3.4)$$

この確率分布は、次に現れるトークンの予測に利用される。

教師ありファインチューニング

教師ありファインチューニングでは、ラベル付きデータセット $C = \{(x, y)\}$ を使用する。ここで、 C はタスクに関連するテキスト（入力トークン列 $x = \{x_1, x_2, \dots, x_m\}$ ）と、それに対応する正解ラベル y のペアで構成されている。

モデルは、入力トークン列 $x_1, x_2, \dots, x_m$ を処理して、最終的な Transformer ブロックの出力ベクトル $\mathbf{h}_l^m$ を生成を行う。このベクトルは、入力テキスト全体の特徴を表現している。次に、この出力ベクトルにタスク特化の線形出力層（重み $\mathbf{W}_y$ ）を適用して、ラベル y の予測確率を計算する。この処理は以下の式で表される。

$$P(y \mid x_1, \dots, x_m) = \text{softmax}(\mathbf{h}_l^m \mathbf{W}_y) \quad (3.5)$$

ここで、 $P(y \mid x_1, \dots, x_m)$ は、モデルが入力トークン列 $x_1, \dots, x_m$ を基にラベル y を予測する確率を示す。

モデルが正しいラベルを高い確率で予測できるようにするために、ラベル付きデータセット C を用いて以下の目的関数を最大化する。

$$L_2(C) = \sum_{(x,y) \in C} \log P(y \mid x_1, \dots, x_m) \quad (3.6)$$

この式は、データセット内のすべてのペア (x, y) について、正しいラベル y を予測する確率の対数を合計するものである。モデルは、この目的関数を最大化するように学習を行う。これが、教師ありファインチューニングの基本的な流れである。

また、学習したモデルの精度をさらに上げる方法として、プロンプトエンジニアリングが挙げられる。プロンプトエンジニアリングは、AI の自然言語処理モデルが最適な結果を出せるように、効果的な質問や指示を設計・最適化する技術である（図 3.3 参照）。タスクをプロンプト形式のデータセットに変換し、学習を通じてモデルの性能を向上させる。プロンプトエンジニアは、適切なプロンプト設計を通じて、AI が意図した応答や文章を生成できるようにする。プロンプトには図 3.4 のような種類があり、例えば以下のような方法があげられる [41]。

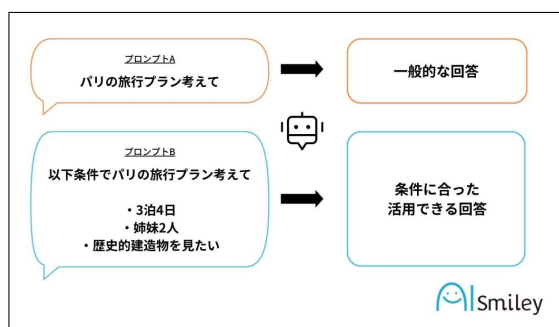


図 3.3: プロンプトエンジニアリングの役割 [41]

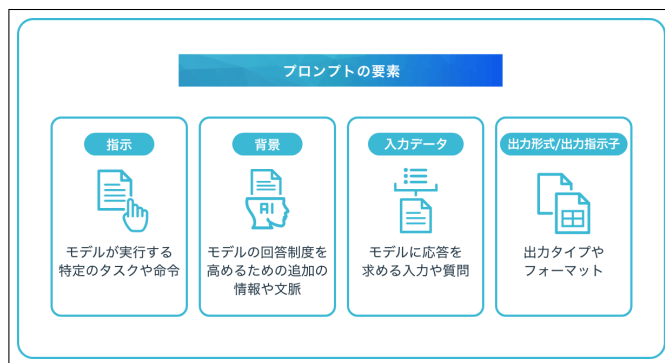


図 3.4: プロンプトの種類 [41]

- **Zero-shot Prompting**

例やデモンストレーションを一切提供せずに、直接質問や指示を与えるプロンプト。大量データでトレーニングされたモデルには有効だが、複雑なタスクでは精度が低下する可能性がある。

- **Few-shot Prompting**

複数の例やデモを提示して、モデルに質問と回答のパターンを学習させる方法。例の数が多いほど、適切な回答が得られる可能性が高まる。

- **Chain-of-Thought Prompting**

複雑な問題解決に必要な段階的な推論を促し、論理的な思考をサポート、数学の証明や論理パズルなど、複雑なタスクに有効。

- **Zero-shot CoT**

CoT の技法を Zero-shot で適用し、「ステップに分けて考えてください」といった指示で推論プロセスを促進。事前例がなくても、推論過程の記述により精度が向上。

- **Self-Consistency**

複数の推論パターンを生成し、その中から最も整合性の高い回答を選択。安定した正確な回答が得られやすくなる。

- **Generate Knowledge Prompting**

問題解決に役立つ知識や情報をプロンプト内に組み込み、より正確な出力を導く手法。事前知識の活用が重要なタスクに有効。

- **ReAct**

「推論」と「行動」を交互に行い、情報収集と実行を繰り返して問題解決を図る手法。行動を伴うタスクや探索型問題に強い。

§ 3.3 問診と共感力

1.1 節でも述べたように、医療の問診において共感力は非常に重要な要素である。そこで、本節では、医療における共感の定義、共感力向上の取り組み、共感力の測定方法について

説明する。

まず、医療分野において共感とは、観察者が視覚的および聴覚的な手がかり（例：表情、声のトーン、姿勢など）を基に対象者の感情状態を特定し、一時的にその感情を体験する能力を指す [42]。この際、その感情がどのような背景や状況から生じたものかを深く理解し、適切なコミュニケーションを通じてその理解を対象者に伝達することが求められる。ここで日本語の「共感」に対応する英語表現である sympathy と empathy の違いを考えると、Sympathy は感情の「共鳴現象」を指し、受け手が他者の感情に共鳴する受動的な反応を表します。一方、empathy は「感情移入」を指し、受け手が相手の感情を能動的に理解しようとする過程を意味する [43]。この違いを踏まえると、共感とは単なる感情移入や同情とは異なり、観察者が対象者の感情を的確に把握し、それに基づいて適切な応答を行うことで信頼関係を構築し、対象者に安心感を与える役割を果たすことがわかる。

共感には主に2つの側面がある。1つは、他者の状況や内面を冷静に理解する認知的共感であり、もう1つは、他者の感情に共鳴する感情的共感である。ここで、感情的共感と認知的共感の特徴について具体的にどのようなものか説明する [43](表 3.2)。

認知的共感の特徴

認知的共感とは、比較的意識的なプロセスを伴うため、ある程度オンとオフを切り替えることが可能である。例えば、他者が罰を受けている場面を目にした際、その人の心理状態を理解し予測することで共感が生じる。しかし、仮にその人が倫理的に不適切な行為をした場合には、「罰を受けるのは当然」と判断し、意識的に共感を抑えることができる。このように、認知的共感は主にトップダウン型の処理として機能すると考えられる。

感情的共感の特徴

感情的共感とは、その名の通り比較的自動的に発生するプロセスを含む。認知的共感のようにオンとオフを切り替えることは一般的に難しく、基本的にはボトムアップ型の処理として機能すると考えられる。

従来の医療現場では、患者に対する冷静で合理的な診療を重視するため、認知的共感が特に重要視されていた。一方で、感情的共感とは判断力を損なう可能性があるとして慎重に扱われ、Detached Concern（認知的共感を持ちながら感情的距離を保つ姿勢）が推奨される傾向にあった。

しかし、近年の研究では、感情的共感が認知的共感を補完する役割を果たし、より正確な患者理解や信頼関係の構築に寄与することが示されている。医療従事者が認知的・感情的な両面からバランスよく共感を示すことで、患者の治療効果や満足度を向上させるという知見が増えている。そのため、現在では Detached Concern に偏らず、認知的共感と感情的共感の調和を図ることが重要視されている。

ポジティブな知見が増える中、医学生や研修医が教育課程で共感性を失いやすいことも報告されている。入学当初は高い共感性を持っていても、臨床実習などを通じて心理的負担が増え、共感性が低下する傾向がある。この課題に対処するため、コミュニケーションスキルトレーニングなどの教育的介入が積極的に導入され、医療従事者の共感性向上が図られている [44]。

表 3.2: それぞれの尺度の違い

項目	認知的共感	感情的共感
特性	意識的で理性的なプロセス	自動的で感情的なプロセス
処理タイプ	トップダウン型	ボトムアップ型
焦点	他者の感情や状況の「理解」	他者の感情や状況への「共鳴」

そこで、医学生や医療従事者の共感力を玉める教育的介入やスキル向上の具体的な方法について説明する [45].

- **コミュニケーションスキル**

患者との信頼関係を構築し、患者中心のケアを実現するために不可欠なスキルとして位置づけられており、ロールプレイや模擬患者を用いたシミュレーション、グループディスカッションがその具体例である。これらは、学生が患者の感情に共感し、それを適切に表現する能力を高めることに貢献する。

- **マインドフルネス**

ストレス軽減や精神的健康の向上を目的としたトレーニングであり、瞑想や呼吸法、反省的学習、マインドフルウォーキングなどが含まれる。これらの活動により、学生が感情的および行動的なバランスを保ちながら、患者への思いやりを持つ姿勢を育成することができる。

- **早期臨床体験**

学生が早期から臨床環境に触れることで、患者の感情や文脈的な状況を深く理解し、患者中心のケアに必要な専門的スキルや価値観を形成する機会を提供する。特に低所得者層を対象とした無料クリニックでの活動は、共感力を向上させる効果が高いとされている。

- **技術を活用した学習**

仮想患者や標準化患者を用いたシミュレーションは、現実的かつ反復可能な学習体験を提供し、共感的なコミュニケーション能力を高める。これにより、安全な環境で学生が実践を重ね、患者中心のケアを学ぶことが可能となる。

- **芸術・文化活動の導入**

映画や文学作品、漫画などを取り入れた教育は、患者を一人の人間として理解する力を育る。これらの活動を通じて、学生は患者の視点を直感的に理解し、患者との人間的なつながりを深めることができる。

- **自己反省と内省**

患者とのつながりを深めるために、学生が自身の行動や態度を振り返る機会を提供する取り組みである。口頭や書面での反省、ブログや物語の共有などが学生の内省能力を高め、共感的なケアを実践する基盤となる。

表 3.3: それぞれの尺度の違い

尺度	特徴	重視するもの
IRI	認知적および感情的共感の多次元評価	認知적共感と感情的共感
HES	認知적共感の評価	認知적共感
EES	他者の感情に対する敏感さや影響を評価	感情的共感
JSE	医療者の患者中心のケアに必要な共感を評価	医療者に必要な認知적共感と感情的共感

このように、様々な共感力向上の取り組みが行われている一方で、それらの効果を正確に把握するためには共感力を測定する必要がある。以下では、共感力を測定する方法についていくつか述べる (表 3.3)。

共感力を測定する手法の1つとして、Interpersonal Reactivity Index (IRI) がある。IRI は、共感を認知적および情動的な多次元的視点から評価するために開発された自己評価尺度であり、

この尺度は、共感が単一の特性ではなく、複数の要素が絡み合った複雑な構造を持つと考え、それを4つの下位尺度で構成されている。1つ目は「Perspective Taking (PT)」であり、これは他者の視点を認識し、それを理解する能力を評価する認知적共感の中心的な要素であり、2つ目が「Fantasy (FS)」であり、これは物語の登場人物や架空の状況に感情移入する力を測定する。さらに、3つ目は「Empathic Concern (EC)」であり、これは他者の困難に対して同情や気遣いといった共感的感情を持つ情動的側面を強調し、4つ目は「Personal Distress (PD)」であり、これは他者の苦しみを目撃した際に自分自身がどれほど苦痛を感じるかを評価する。

IRI は、認知적および情動的な共感を包括的に評価できる多次元的アプローチであり、その内部一貫性や妥当性の高さから、医療、心理学、教育、カウンセリング、ビジネスなど、さまざまな分野で広く使用されている。自己評価形式の28項目で構成され、リッカート尺度を用いて回答するため、信頼性の高い測定が可能である。その利点として、共感の認知적および情動的側面を詳細に把握できる点や、研究や教育プログラムにおける応用可能性の高さが挙げられる。一方で、文化的背景や言語による適応の必要性、回答が社会的望ましさバイアスに影響を受ける可能性など、いくつかの課題も存在する [46]。

二つ目は、Hogan Empathy Scale (HES) である。HES は、共感を「他者の立場や感情を知覚的または想像的に理解する能力」と定義し、主に認知적共感に焦点を当てており、Minnesota Multiphasic Personality Inventory (MMPI) や California Psychological Inventory (CPI) から関連する質問を抽出し、64項目の二者択一形式で構成されている。妥当性の検証では、臨床心理士やカウンセラーなど他者理解が求められる職業の人々が高いスコアを示し、尺度の有効性が確認された。HES は特に認知적視点取得能力を重視し、医療や教育、カウンセリング、ビジネスなど幅広い分野で活用されている。しかし、情動的共感を軽視する点や文化的適応の必要性、自己申告形式によるバイアスの可能性が課題として指摘されている [47]。

三つ目は、Emotional Empathy Scale (EES) である。EES は、感情的共感を測定するために設計されている。この尺度は、他者の感情体験に対して感情的に反応する能力を評価することを目的としており、感情的共感を認知적共感から明確に区別している。

表 3.4: EES の項目の一部

No.	Question	Sign
1	It makes me sad to see a lonely stranger in a group	+
2	People make too much of the feelings and sensitivity of animals	-
3	I often find public displays of affection annoying	-
4	I am annoyed by unhappy people who are just sorry for themselves	-
5	I become nervous if others around me seem to be nervous	+
6	I find it silly for people to cry out of happiness	-
7	I tend to get emotionally involved with a friend's problems	+
8	Sometimes the words of a love song can move me deeply	+
9	I tend to lose control when I am bringing bad news to people	+
10	The people around me have a great influence on my moods	+

EES は 33 項目から構成されており、感情の伝染への感受性や見知らぬ他者の感情理解、他者のポジティブ・ネガティブな体験に対する感情的反応、同情的傾向、問題を抱える人々との関わりへの意欲など、感情的共感のさまざまな側面を評価する。各項目はリッカート尺度で測定されており、具体的には、表 3.4 のような質問がある。EES は心理学的研究やカウンセリング、医療現場、人間関係の分析などで広く活用されており、感情的共感の多次元的な評価を行うことができる [48]。

四つ目は、Jefferson Empathy Scale(JSE) である。JSE は、医師や医学生の共感性を測定するために開発された尺度であり、患者中心のケアにおける共感を、医療者の態度、行動、認知的理解の観点から評価する。この尺度は、患者の視点を理解する能力 (Perspective Taking)、患者への配慮を示す行動 (Compassionate Care)、患者の個別性の認識 (Standing in the Patient's Shoes) の 3 つの主要な側面を評価するために設計されており、リッカート尺度で回答する。開発にあたっては、医療現場に特化した項目が設計されており、医療教育における共感力育成プログラムの効果測定、臨床実践での医療従事者の共感性モニタリング、医療現場での共感の役割を解明する研究など、多岐にわたる分野で使用されている [49]。

また、JSE は、対象者に応じて以下の 3 つのバージョンが存在する。1 つ目は、JSE-S という医学生向けのバージョンであり、これは医学生の共感性を評価するために設計されている。2 つ目は、医療専門職向けの JSE-HP バージョンであり、これは医師や看護師、歯科医師、薬剤師、臨床心理士など、患者ケアに携わる医療専門職の共感性を測定するためのものである。3 つ目は、医療専門職学生向けの JSE-P で、これは医学生以外の医療専門職を目指す学生の共感性を評価するために開発されたものである。

提案手法

§ 4.1 漢方医学的問診とバーチャル患者

本研究では、漢方医学的問診ができるバーチャル患者の開発を行う。ここでバーチャル患者とは、医療教育やトレーニングのために設計されたコンピュータのシミュレーション患者のことを指す。この技術は、医学生や医療従事者が患者と対話し、診断や治療計画を実践的に学ぶ場を提供する目的で開発されている。バーチャル患者は、実際の患者を模倣し、対話を通じて症状や感情を表現するだけでなく、感情的な反応や身体的な特徴を動的に変化させることが可能である。これにより、医療シナリオを再現しながら、学習者に安全な環境での繰り返し練習を行えるようになる。今回、開発を行うバーチャル患者は、音声で対話を行い、漢方医学的問診のシナリオに沿った回答を行うものである。そこで、本節では、本研究の対話型システム、漢方医学的問診シナリオの作成、漢方医学的問診を行うための生成 AI の学習、対話型システムと生成 AI の統合について述べ、どのようにバーチャル患者を開発したのか説明する。

本研究の対話型システム

まず、本研究の対話型システムについて説明する。対話型システムは、3.1 節でも述べたように、音声認識技術とテキスト合成音声技術が用いられている。本研究では、音声認識技術に Google Cloud Speech-to-Text、合成音声技術に VOICEVOX を用いる。

Google Cloud Speech-to-Text

Google Cloud が提供する音声認識 API であり、音声データをテキストに変換するサービスである。このサービスは、リアルタイムでの音声認識や、録音済みの音声ファイルの文字起こしを高精度で実行できるのが特徴である。また、Google の AI 技術を活用し、多言語対応やカスタマイズ機能が充実しているため、幅広い用途で活用されている [50]。

VOICEVOX

無料で利用できる高品質なテキスト音声合成ソフトウェアであり、音声合成技術を用いて、入力したテキストを自然な音声に変換できるのが特徴である。オープンソースで開発されており、商用・非商用を問わず無料で利用することができる。また、発話の速度や抑揚を調整することができるため、用途に応じた表現が可能である [51]。

表 4.1: 問診シナリオの一部

No.	学生	バーチャル患者
1	食欲はありますか？	あまり食欲はないですね
2	下痢はしやすいですか？	緩い便のことが多くて、下痢することもあります。
3	暑がりですか？寒がりですか？	寒がりです。
4	熱いお風呂とぬるいお風呂のどちらに入られますか？	熱い方が好きです。
5	疲れやすいですか？	疲れやすいです。
6	冷房は大丈夫ですか？	苦手です。
7	他に困っている症状はありますか？	疲れやすいです。
8	夢はよく見ますか？	時々は見ますが、そんなに多くはないです。
9	お腹が張ることはありますか？	そうですね。食後に張るかな。
10	手足は冷えますか。	はい。手足は冷えやすいです。特に冬は冷えるので、手袋したり靴下2枚履いたりします。カイロも使いますね。

漢方医学的問診シナリオの作成

次に漢方医学的問診シナリオの作成について述べる。まず、問診シナリオとは、特定の主訴や症状もつ患者に対して、医学生が漢方医学的病態を具体的に質問し、患者がそれに回答する状況を設定したものである。

本研究では、漢方薬の中でも処方頻度が高い「人参湯(にんじんとう)」を題材にした問診シナリオを作成した。人参湯は、胃腸炎や食欲不振の治療に用いられる漢方薬である。シナリオの構築に際しては、実際に人参湯を処方された患者の漢方医学的問診を参考にし、胃部不快感および食欲不振を主訴とする60代後半の男性患者をモデルとした問診シナリオを作成した。このシナリオは、4名の漢方専門医と複数回のディスカッションを経て改良を重ねたものである。表4.1がシナリオの一部である。このシナリオを理解するにあつた、漢方医学が属する東洋医学と、西洋医学の違いを再度明確にしておく必要がある。

まず東洋医学は、東洋医学（ここでは漢方医学や鍼灸医学など日本の伝統医学を指す）は、「個の医学」として知られ、個々の体質や特徴を重視する。その基盤には、心と身体が一体であるという「心身一如」の考え方があり、特定の器官に焦点を当てるのではなく、身体全体の調和を図る全人的医療を目指している。このように東洋医学は、心身一元論に基づいて発展してきた医学体系である。診察では患者の「主訴」に偏ることなく、切診を中心に行い、「人」の全体的な調和を観察して異常を診断する。つまり、東洋医学的診察では患者の自覚症状に頼らず、その状態を包括的に把握することが可能である。

次に西洋医学は、自然科学を基盤として普遍性、客観性、論理性を重視する医学体系であり、主観を排除するために「主体」と「客体」を分離する心身二元論の立場を取る。このため、西洋医学における診療の手順は、「主訴」→「診察」→「診断」→「治療」という流れで進行する。診察では、主訴が診療の焦点となり、その部位を中心にレントゲンやMRI検査などが行われる。検査結果は客観性を高めるために数値化や計量化が進められ、その技術の進歩により精度は向上している。しかし、このような要素還元的なアプローチは、診断が細分化されすぎることによって主訴に偏り、疾病の全体像を把握しにくくなるというリスクもある。[52]。

```

{"messages": [{"role": "system", "content": "あなたは67歳の男性患者で、名前は杉谷富介です。医療問診の場における患者として、医師の質問に対して適切に回答してください。"}, {"role": "user", "content": "食欲はありますか?"}, {"role": "assistant", "content": "あまり食欲はないですね。"}]}
{"messages": [{"role": "system", "content": "あなたは67歳の男性患者で、名前は杉谷富介です。医療問診の場における患者として、医師の質問に対して適切に回答してください。"}, {"role": "user", "content": "冷房は大丈夫ですか?"}, {"role": "assistant", "content": "苦手です。"}]}
{"messages": [{"role": "system", "content": "あなたは67歳の男性患者で、名前は杉谷富介です。医療問診の場における患者として、医師の質問に対して適切に回答してください。"}, {"role": "user", "content": "疲れやすいですか?"}, {"role": "assistant", "content": "疲れやすいです。"}]}
{"messages": [{"role": "system", "content": "あなたは67歳の男性患者で、名前は杉谷富介です。医療問診の場における患者として、医師の質問に対して適切に回答してください。"}, {"role": "user", "content": "下痢はしやすいですか?"}, {"role": "assistant", "content": "緩い便のことが多くて、下痢することもあります。"}]}

```

図 4.1: トレーニングデータの内容

漢方医学的問診を行うための生成 AI の学習

次に漢方医学的問診を行うための生成 AI の学習方法について述べる。本研究では、ChatGPT という生成 AI を用いる。

ChatGPT

ChatGPT は、OpenAI が開発した生成 AI であり、ユーザーの入力に応じて自然な文章を生成することができるモデルである。Generative Pre-trained Transformer という技術を基盤にしており、大量のテキストデータを学習することで、多様な質問や会話に対応できる。ChatGPT は、日常会話のほか、プログラミングの支援、文章の作成、翻訳、要約、アイデア出しなど、幅広い用途に活用できる。特に、プロンプトによって出力内容を調整することで、カスタマイズ性の高い応答を提供できる。さらに、API を利用することでアプリケーションやサービスに組み込むことも可能である [53]。

この ChatGPT に今回作成した問診シナリオを学習させることで、漢方医学的問診を行えるようにする。学習には3.2節で説明した、ファインチューニングを行う。そこで、ChatGPT のファインチューニングの流れについて説明する。

最初のステップとして、ファインチューニングに使用するデータセットを準備する。このデータセットは、モデルに学習させたい具体的なタスクや応答スタイルを示す例を含む必要があります。ChatGPT のファインチューニングでは、データは会話形式で提供する必要があります。以下の3つの要素を含む構造を取る。まず、「システムの指示 (system)」は、モデルに与える背景や役割を定義する要素である。次に、「ユーザーの入力 (user)」は、モデルに対してユーザーが投げかける質問や発言を示す。これは実際の会話でモデルがどのように質問を受けるかを再現する部分である。最後に、「アシスタントの応答 (assistant)」は、モデルがどのように応答するべきかを示すものであり、期待する出力の例を提供する。この部分では、モデルに期待される応答のスタイルや内容が具体的に記述されている必要が

表 4.2: モデルの比較

モデル	推論のテキスト長	トレーニング時のテキスト長	特徴
gpt-4o-2024-08-06	128,000トークン	65,536トークン	高性能モデル、長文の処理に優れる
gpt-4o-mini-2024-07-18	128,000トークン	65,536トークン	低コスト版、計算効率を重視
gpt-3.5-turbo-0125	16,385トークン	16,385トークン	GPT-3.5の改良版
gpt-3.5-turbo-1106	16,385トークン	16,385トークン	さらなる最適化が施されたバージョン
gpt-3.5-turbo-0613	16,385トークン	4,096トークン	コンテキスト長が短い、軽量で効率的

ある。本研究で、用意したトレーニングデータの一部を図 4.1 に示す。user と assistant には、本研究で作成した問診シナリオの会話を記述し、system には、問診シナリオのモデルになる人物の一般的な内容を記入した。

データセットを準備したら、次に OpenAI API を通じてデータをアップロードを行う。アップロードは Files エンドポイントを使用する。データのアップロードが完了した後は、ファインチューニングジョブを作成し、モデルのトレーニングを行う。このトレーニングを行うには、いくつかの情報を指定する必要がある。

まず、使用するベースモデルを選択する。例えば、gpt-4o-mini-2024-07-18 などの適切なモデルを指定を行う。次に、事前にアップロードしたトレーニングデータをトレーニングファイルとして指定する。ベースモデルには、表 4.2 のように様々なものがあるが、本研究では、一番性能の高い gpt-4o-2024-08-06 を用いた。

これらの指定を行った後、ファインチューニングジョブを開始し、モデルの調整を行う。トレーニングが完了すると、ファインチューニングされた新しいモデルが生成され、指定されたタスクやスタイルに適した応答を生成できるようになる [54]。

対話型システムと生成 AI の統合

まず、音声認識技術を用いてユーザーが行う問診の音声をリアルタイムでテキストに変換する。変換されたテキストは、漢方医学的な問診シナリオを学習した ChatGPT に送られ、適切な応答が生成される。この生成された応答は、テキスト音声合成システムに送られ、音声として出力される。この一連の流れにより、漢方医学の問診プロセスを適切に再現できるバーチャル患者システムが構築される。これにより、医学生は実際の患者と対話するような体験を通じて、問診スキルや診断能力を効率的に学ぶことが可能となる。

§ 4.2 感情推定と表情の選択

本研究では、バーチャル患者を人型の 3D モデル (3D キャラクター) に適応し、リアリティのあるシステムの作成を行った。また、感情推定を行い感情に合わせた表情の選択を行うことで、感情を反映したシステムを作成し共感力の向上を計る。

表 4.3: 感情推定の種類

手法の種類	概要	方法
テキストベース	テキストデータから感情を分析	・ルールベース: 感情辞書を用いる ・機械学習: ロジスティック回帰など ・ディープラーニング: BERT, GPTなど文脈を考慮したモデル
音声ベース	音声データから感情を推定. 声のトーン, 速さなどを活用.	・音響特徴抽出: MFCCなど ・機械学習: ランダムフォレストなど ・ディープラーニング: CNN, LSTMなど
表情ベース	顔の表情から感情を推定. 表情筋の動きや顔の特徴点を分析.	・顔特徴点抽出: ランドマーク検出 ・機械学習: OpenCV ・ディープラーニング: CNNで画像から直接感情を推定
生体情報ベース	心拍数, 脳波, 皮膚電位などの生体データを用いて感情を推定.	・生体信号解析: 心拍数, 皮膚電気活動, 脳波を解析

3D キャラクターの作成

本研究では, 漢方医学的問診ができる対話型システムに, 3D キャラクターを適用した. 3D キャラクターとは, 3 次元空間上で作成された立体的なキャラクターのことを指す. 3D キャラクターはコンピュータグラフィックス技術を使って制作され, 主にアニメーションやゲームといったデジタルコンテンツで広く活用されている. 2D のイラストやアニメーションと比較すると, 3D キャラクターはよりリアルな表現が可能で, 立体的な形状や自然な動きを忠実に再現できる点が特徴である [55]. 本研究では, CharacterCreator4 という 3D キャラクター作成ソフトを用いて, 3D キャラクターの作成を行った.

CharacterCreator4

CharacterCreator4 とは, Reallusion 社が提供する高度な 3D キャラクターデザインツールである. このソフトウェアは, リアルな 3D キャラクターやアニメ調のキャラクターを効率的に作成, 編集, アニメーション化するための機能を備えている. また, 身体のプロポーション, 顔のパーツ, 肌のテクスチャ, 衣装などを細部まで調整可能であり, キャラクターに表情を付けて, 感情豊かな表現を可能にすることができる [56].

ここで, 3D キャラクターには喜び (Happiness), 悲しみ (Sadness), 怒り (Anger), 恐れ (Fear), 嫌悪 (Disgust), 驚き (Surprise) の 6 つの表情を作成した. この 6 つの感情は, エクマンの基本感情と呼ばれ, 人間が普遍的に持つとされる感情のことである. 彼の研究は, 文化や環境に関係なく, 人間が特定の基本感情を共通して持ち, それらが顔の表情として現れることを明らかにされた [57]. そこで, それぞれの感情の表情はどのようなものなのか説明する. まず, Happiness の表情は, 口角を引き上げることで形成され, この際に目の周りにしわができ頬が引き上げられ笑顔の表情となる. 次に, Sadness の表情は, 眉の中央が持ち上げられることで特徴づけられる. これにより, 目尻が下がり, 全体的に沈んだ印象

この会話を読み、エクマンの6つの基本感情 (Happiness、Sadness、Anger、Fear、Disgust、Surprise) をそれぞれ0~10のスケールで評価してください。0はその感情が全く表れていない場合、10はその感情が非常に強く表れている場合を示します。

```
{
  "感情推定結果": {
    "Happiness": 0~10,
    "Sadness": 0~10,
    "Anger": 0~10,
    "Fear": 0~10,
    "Disgust": 0~10,
    "Surprise": 0~10
  }
}
```

図 4.2: 感情推定のプロンプト

を与える。また、口角も下げられ、ときには唇がわずかに震えることもある。次に、Angerの表情は、眉間にしわを寄せ、眉を中央に引き下げることによって特徴づけられる。また、目は細められ怒った目つきとなり、唇は強く閉じられるか、歯を食いしばる形になる。また、顎が突き出されことも多い。次に、Fearの表情は、目を大きく開く眉を上げて内側に寄せることによって特徴づけられる。次に、Disgustの表情は、鼻にしわを寄せ上唇が持ち上げられることによって特徴づけられる。また、口角が下げられる、あるいは唇がすぼめられることで、不快感や嫌悪感を表現す。最後に驚き Surpriseの表情は、眉が高く持ち上げられる、目が大きく開かれることによって特徴づけられる。さらに、口が開かれることで、瞬間的に驚いた反応を示す [57]。このような、感情に対する表情が特徴づけられているため、これらを参考にし表情の作成を行う。

また、3D キャラクターに感情に対する視線の動きも取り入れることにした。一般的に、人間はネガティブな感情を抱いているときに視線を下にそらす傾向があるとされている。この動作は心理学や非言語コミュニケーションの分野で観察されており、感情状態や心理的な内省を反映するものと考えられている。さらに、顔の表情の認識に関しては、西洋人が口元に注目する傾向があるのに対し、日本人は目に注目する傾向があることが知られている。また、漢方専門医は患者の目を注意深く観察し、そこから感情を読み取って共感を示すことが多いとされる。これらの知見に基づき、バーチャル患者の目の動きをリアルに再現することで、医学生が患者の感情をよりの確に察知し共感する能力を向上させることができるのではないかと考えられる。

感情推定

感情推定には、表 4.3 のように様々な方法がある。本研究では、テキストベースの手法である ChatGPT を用いた方法で行う。なぜ ChatGPT を使うのかというと、まず、細かく多次元的な感情を正確に抽出できる能力を持つことが、先行研究により示されているからである。さらに、ChatGPT は Valence, Arousal, Dominance といった感情次元を理解し、それらを基に感情を推定する能力を持つとされている。この多次元的な理解は、エクマンの

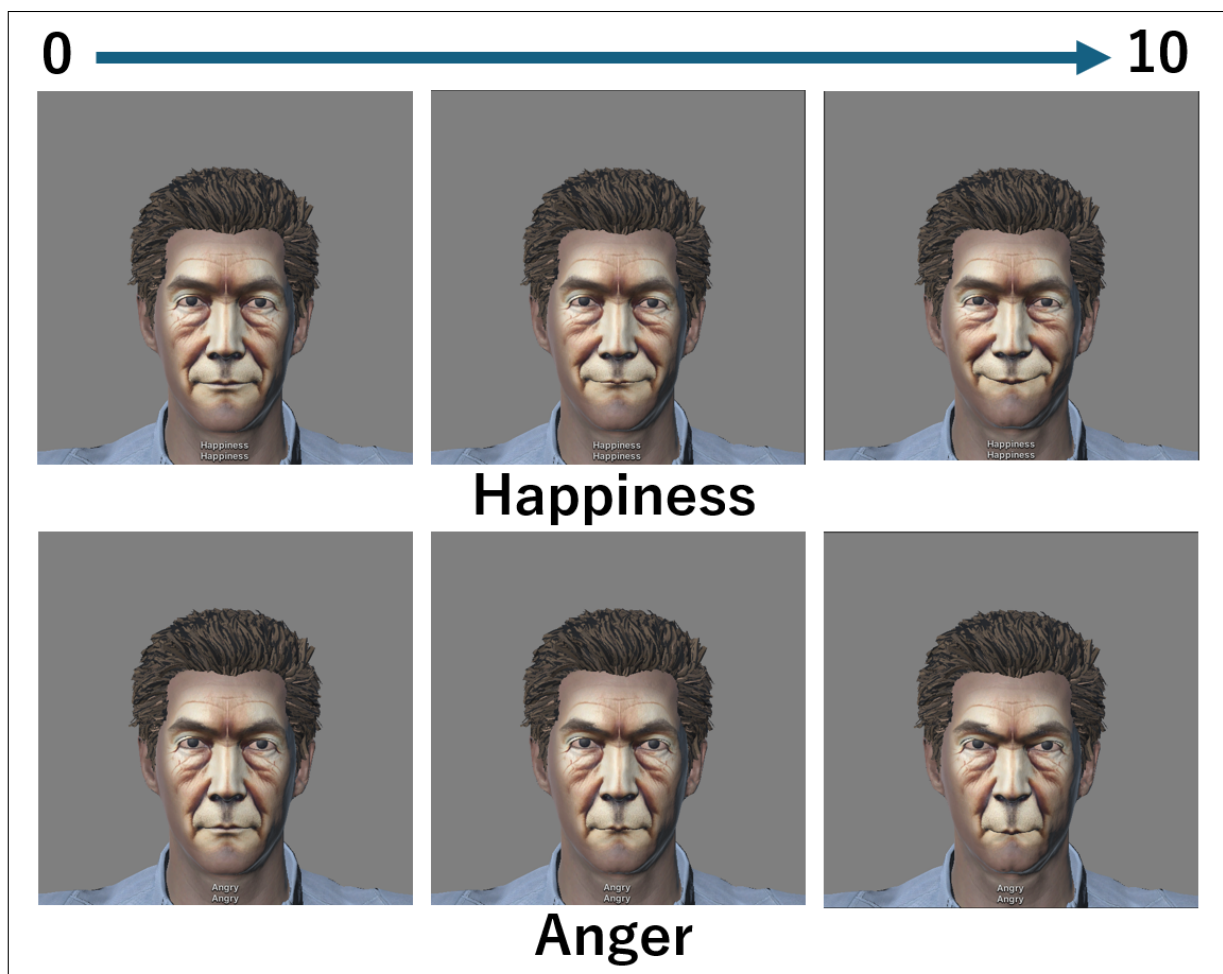


図 4.3: 感情の強度に対する表情

6つの感情を明確に分類する際の基盤となり、より精度の高い感情推定が可能となる。加えて、ChatGPTは、適切なプロンプトを用いることで感情の誘発や推定が可能であることが示されている [58]。本研究では、この特性を活用し、感情推定における効果的なプロンプト設計を行う。

具体的な方法としては、ユーザー側の入力と、4.1節で説明した問診シナリオを学習させたモデルの出力を、エクマンの基本感情である6つの感情の推定を行うようにプロンプトエンジニアリングしたChatGPTに送り、感情推定を行う。具体的なプロンプトを図4.2に示し、どうしてこのように設計したのか述べる。

まず、感情を「0～10」のスケールで評価する形式にした。これにより、感情の強度をより細かく定量的に評価できるようになる。0は感情が全く表れていない状態であり10は非常に強く表れている状態を表す。このように明確なスケールを提示することで、モデルが一貫性のある評価が行えるようになる。

次に、出力結果の形式を明示し、各感情カテゴリを「感情名: スコア」の形でリスト化して出力するように指示した。この指定により、得られる結果が一貫性を持ち、解析や評価を容易にすることができる。またそれぞれの感情カテゴリを、Happiness, Sadness, Anger, Fear, Disgust, Surpriseの英語で指定した。これは、モデルの性能は使用するデータセッ

トや感情ラベルの表現に大きく依存しており、結果が一貫しない場合がある [59] ためである。このような曖昧さを最小限に抑えるために、統一された英語表現を採用し、モデルが感情カテゴリを正確に解釈できるようにした。

感情と表情の一致

本研究では、推定された感情の種類と強度に応じて、適切な表情を適用することとする。これにより、より自然で直感的な感情表現を行えるようになる。また、複数の感情が同時に推定された場合は、それぞれの感情に対応する表情を統合し、複合的な表情を生成することにする。これにより、例えば「Anger」と「Happiness」が同時に高いスコアで推定された場合には、驚きの要素を含みつつも、明るい表情を維持することができる。さらに、感情の強度に基づいて表情の変化を段階的に調整することで、より繊細でリアルな感情表現を実現を行う。この時、感情値に対する表情は以下ようになる。

- **スコアが0～3の場合**

この範囲では、感情の表出がほとんど見られないため、通常時の表情に近い表情となる。特定の感情が顕著に表れていないときや、静かな状態にある場合などに適用される。

- **スコアが4～6の場合**

感情が中程度に表れている状態とし、軽い感情表現が行われる。具体的には、少し微笑む、軽い驚きを見せるなど、感情を抑えた控えめな表情を選択する。

- **スコアが7～10の場合**

感情が強く表れている状態とみなし、明確な感情表現を採用する。例えば、大きく笑う、驚きで目を見開くなど、感情を強く示す表情がこれに該当する。

このように表情を選択する理由は、エクマンによる研究で、感情の強度が表情に大きな影響を与えることが明らかになっているからである [60]。感情が強くなるほど顔の筋肉の動きが顕著になり、表情がよりはっきりと現れるとされている。感情の強度を考慮した表情は、図 4.3 のようになった。また、人間はネガティブな感情を抱いているときに視線を下にそらす傾向があるとされているため、ネガティブな感情が推定された場合は、視線を下にそらすことにする。エクマンの基本感情は一般的に次のように分類される。Happiness はポジティブな感情であり、Anger, Sadness, Fear, Disgust はネガティブな感情と考えられる。Surprise については、ポジティブにもネガティブにも偏らない中立的な感情とみなされることが多い。

そこで、ネガティブな感情と考えられている、Anger, Sadness, Fear, Disgust のスコアを合算し、その合計が5以上に達した場合には、視線を下に向ける動作を加えることにする。これにより、人間らしい感情表現を実現することができると考える。このように、感情推定の結果に基づいて段階的な表情変化を適用し、視線の動きと組み合わせることで、より豊かで自然なコミュニケーションを実現することを目指す。

表 4.4: JSE の一部 [61]

下記のそれぞれの文章についてあなたがどの程度同意するか、あるいは同意しないかを表示してください。
 その際、適切な評価指数（後述のスケール参照）を選び右欄の○を塗りつぶしてください。
 以下の 7 点のスケールを使用して下さい（点が高いほど強く同意することを示します）。それぞれの設問において塗りつ
 ぶすのは一つだけです。1（全く同意しない）2-3-4-5-6-7（全くその通りである）

	1	2	3	4	5	6	7
1 担当患者及び家族がどのように感じているかということを理解することは、内科的、外科的な治療に影響を及ぼさない。	☐	☐	☐	☐	☐	☐	☐
2 私が担当患者の気持ちを理解したら、担当患者はより快適と感じる。	☐	☐	☐	☐	☐	☐	☐
3 私にとって担当患者の視点に立って物事を見ることは難しい。	☐	☐	☐	☐	☐	☐	☐
4 担当患者のボディランゲージを理解することは、治療者-患者関係において言葉によるコミュニケーションと同じくらい重要であると思う。	☐	☐	☐	☐	☐	☐	☐
5 私はユーモアのセンスがあり、それはより良い臨床的結果をもたらすと考える。	☐	☐	☐	☐	☐	☐	☐
6 人はそれぞれに異なっているので、自分の患者の視点で物事を見ることは困難である。	☐	☐	☐	☐	☐	☐	☐
7 私は病歴聴取の際や身体的健康状態について尋ねる際に、患者の感情に留意しないようにしている。	☐	☐	☐	☐	☐	☐	☐

共感力を評価する手法

本研究では、共感力が向上したかを検証するために、共感力を評価する必要がある。3.3 節でも述べたように、共感力を評価する手法はいろいろあるが、本研究では JSE を用いることとした。理由としては、JSE は医療分野において特に使用される共感力を定量的に評価するための指標として広く認知されているからである。また、この尺度は医療従事者や医療教育に特化した設計になっており、多くの研究でその信頼性と妥当性が実証されている。さらに、JSE は表 4.4 のように日本語版も開発されており、その信頼性と妥当性が検証されている。日本語版 JSE では、「他者の視点に立つこと（perspective taking）」や「思いやりのあるケア（compassionate care）」など、米国やメキシコで抽出された構成要素と同様の結果が得られており、日本の医療分野における適用性が確認されている [61]。このように、日本語版 JSE は本研究において、日本語環境で医学生の共感力を適切に評価する最も適しているツールであると考えた。

§ 4.3 提案システムのアルゴリズム

本研究では、4.1 節と 4.2 節で述べたシステムを組み合わせることで、漢方医学的問診と共感力向上のためのバーチャル患者モデルの開発を行う。そこで、システムの統合とシステム全体の流れについて説明する。

システムの統合

本研究では、Unity および ChatdollKit を用いて、全体の統合を行う。

Unity

Unity は、ゲームエンジンの一つであり、2D および 3D のゲーム開発、シミュレーション、仮想現実・拡張現実アプリなど、多岐にわたる開発に利用されている。また、

表 4.5: ChatGPT の設定項目

パラメーター	内容
model	使用するGPTモデル
temperature	応答のランダム性
max_tokens	出力の最大トークン数

表 4.6: VOICEVOX の設定項目

パラメーター	内容
Speaker	話者
Speed Scale	話す速度
Pitch Scale	音の高さ
Intonation Scale	抑揚

Unity は、マルチプラットフォーム対応に優れており、一度開発したゲームやアプリケーションを Windows, Android など様々なプラットフォームに簡単に展開できることが特徴である [62].

ChatdollKit

ChatdollKit は、Unity 上で対話型 AI を構築・統合するためことができるのフレームワークである。音声認識、自然言語処理、音声合成、アニメーション制御などを統合し、インタラクティブなシステムの作成を行うことができる [63].

これにより、漢方医学的問診に必要な複数の要素を効率的に連携させることが可能になった。

システムの設定

システムを動かす前に表 4.5 や表 4.6 のような設定する項目があるので、どのようなものがあるのか述べる。まず、ChatGPT の設定項目では、使用する GPT モデル (model) を指定し、応答のランダム性 (temperature) を調整することで、より創造的または一貫性のある応答を得ることができる。また、出力の最大トークン数 (max_tokens) を設定することで、応答の長さを制御することが可能となる。今回は、回答生成の ChatGPT は、model を gpt-4o-2024-08-06, temperature を 1, max_tokens を 100 とした。感情推定の ChatGPT は、model を gpt-4o, temperature を 0.5, max_tokens を 50 とした。

一方、VOICEVOX の設定項目では、音声合成の詳細な調整を行う。Speaker では使用する話者の選択を行い、Speed Scale により話す速度を変更できる。さらに、Pitch Scale によって音の高さを調整し、Intonation Scale を用いて抑揚の強さを設定することができる。これらのパラメータを適切に設定することで、より自然な対話と音声合成を実現することが可能となる。今回は、Speaker を 13, Speed Scale を 0.8, Pitch Scale を -0.15, Intonation Scale を 1.2 とすることで、今回の設定の人物である高齢の男性に近い音声になるようにした。

システムの流れ

本システムを使用した際の流れは、図 4.4 のようになっており、以下の手順に基づき詳しく説明する。

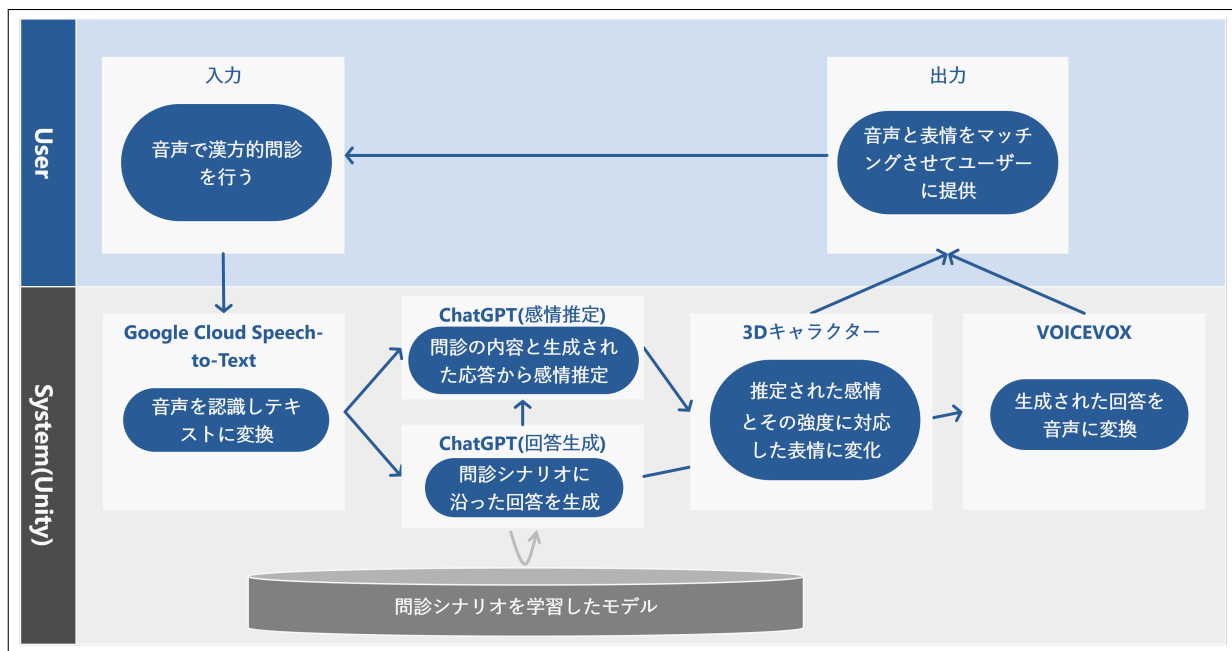


図 4.4: 提案システムの流れ

Step 1: 漢方医学的問診の実施

まず、ユーザが本研究で作成した、バーチャル患者に対し、漢方医学的問診を行う。例えば、ユーザが「食欲はありますか」と音声で質問を行ったとする。この音声は、音声認識技術の Google Cloud Speech-to-Text を用いてテキストデータに変換される。その後、このテキストデータを問診シナリオを学習した ChatGPT に送信し、適切な回答を生成する。

例えば、ChatGPT が適切な回答として「あまり食欲はないですね」と回答した場合、この内容が次のステップに引き継がれる。

Step 2: 感情推定の実施

次に、バーチャル患者の感情推定を行う、前のステップで引き継がれたユーザーの質問と前のステップで生成された回答にあたる、バーチャル患者の回答テキストを、感情推定専用の ChatGPT モデルに送信する。感情推定の結果として、以下のような感情値が出力されるとする。

- Happiness: 0
- Sadness: 3
- Anger: 0
- Fear: 2
- Disgust: 1
- Surprise: 0

この例では、最も高い感情値は Sadness (3) であり、バーチャル患者の感情として悲しみが一番大きかったこととなる。

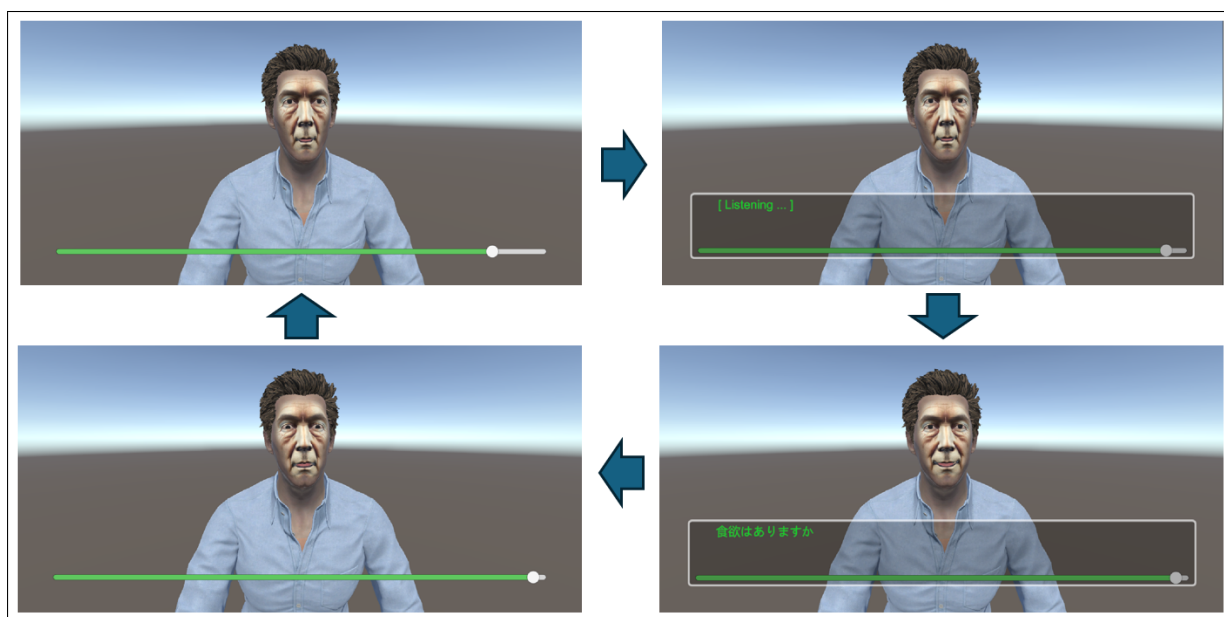


図 4.5: 実際のシステムの推移

Step 3: 表情選択

次に、感情値に基づいてバーチャル患者の表情を選択する。今回の例では、感情は Sadness, Fear, Anger にそれぞれ数値があり、値としては、3, 2, 1であった。この感情値は、本研究で設定した「感情強度に応じた表情基準」に従い、推定された三つの感情を感情の表出がほとんど見られないほどの少しの変化で複合した表情となる。

Step 4: 音声と表情の出力

最後に、Step1 で ChatGPT が生成した「あまり食欲はないですね」を、テキスト音声合成技術の VOICEVOX を用いて音声化する。音声化された応答は、Step3 で選択された表情と同期され、ユーザに出力される。今回の例では、バーチャル患者は通常に近い複合された表情で「あまり食欲はないですね」と音声で返答することになる。

以上が、本研究におけるシステムの一連の流れである。これにより、漢方医学的問診におけるユーザーの質問からバーチャル患者の応答、感情推定、表情変化、音声出力までを統合的に処理し、自然でリアルな対話体験を提供することができる。実際にシステムが動いた時の例を、図 4.5 に示す。

数値実験並びに考察

§ 5.1 数値実験の概要

本研究では、4章で説明した漢方医学的問診と共感力向上のためのバーチャル患者モデルに対して、実際に漢方医学的問診を行い、回答の内容、感情推定の結果、感情に対する表情の測定を行う。また、本システムを使用してもらう前後で共感力の測定を行い、得られた結果から共感力の向上に寄与したのか分析を行い本システムの有効性を示す。

システムの環境

本研究で使用した環境を表 5.1 に示す。まず Unity は、2022.3.33f1 を用いた。このバージョンは、Unity2022LTS 内のバージョンであるため、安定性が高く、長期間サポートされるため、本研究に適していると考えた。次に ChatGPT（応答生成）は、gpt-4o-2024-08-06 を使用した。このバージョンは、最新のバージョンであり応答速度が速く、より高精度な推論が可能であるためである。同様に ChatGPT（感情推定）も最新のバージョンである gpt-4o を使用した。また、Google Cloud Speech-to-Text は v1 を使用し、VOICEVOX は、最新のバージョンである 0.22.3 を使用した。

バーチャル患者の設定

回答生成する生成 AI に学習を行う時に使用したシナリオのモデルになった患者の情報について説明する。名前は杉谷富介であり、67 歳の男性で、身長 165cm、体重 55kg である。現在は無職であり、生活習慣として喫煙はなく、飲酒は嗜む程度である。また、特にアレルギーはない。

主訴として、胃部不快感と下痢の継続が挙げられ、既往歴としては、55 歳頃に高血圧症と診断され、現在もカルシウム拮抗薬（朝 1 回、17mg）を服用している。また、バイタルサインは、血圧 128/86mmHg、脈拍 68 回/分、呼吸数 18 回/分、体温 35.8℃と安定している。

現病歴として、60 歳頃から胃のむかつきやもたれを自覚するようになった。仕事のストレスが原因ではないかと考え、上部消化管内視鏡検査を受けたが、特に異常は見つからなかった。その後、プロトンポンプ阻害薬を服用することで症状は一時的に緩和されたが、65 歳で定年退職した後も胃部不快感が継続し、時折症状が悪化することがあった。さらに、下痢が続くため、近医で整腸剤を処方されたが、効果は一時的であった。今回、家族の勧めもあり、新たな治療法を求めて初めて漢方外来を受診したという設定である。

表 5.1: 使用したシステムの環境

システム	バージョン
Unity	2022.3.33f1
ChatGPT(応答生成)	gpt-4o-2024-08-06
ChatGPT(感情推定)	gpt-4o
Google Cloud Speech-to-Text	v1
VOICEVOX	0.22.3

実験方法

本実験では、被験者に対して JES を用いたアンケートを実施し、本システム使用前後における共感力の変化を評価する。共感力の測定を行うことで、バーチャル患者との対話が被験者の共感力にどのような影響を与えるのかを検証することを目的とする。

まず、実験の開始前に、被験者には事前アンケートとして JES に回答してもらう。これは、システムを使用する前の被験者の共感力のベースラインを測定するために行う。その後、本システムを用いた問診を実施し、問診時間は 10 分間に設定する。問診の進行中、被験者は自由にバーチャル患者と対話してもらう。

問診が終了した後、被験者には再度 JES を用いたアンケートに回答してもらい、システム使用後の共感力の変化を測定する。これにより、バーチャル患者との対話が被験者の共感力に与える影響を定量的に評価することができる。

被験者は 6 名の学生を対象とし、本システムの評価を行う。なお、本システムは音声認識を用いた対話システムであるため、実験は静かな環境で実施し、ノイズの影響を最小限に抑えるよう配慮した。

JES の分析方法

本研究で用いた JES の分析方法について説明する。まず JES は、各項目について 1～7 のスコアで回答を取得しするものである。また、JES には共感力が低いときに高得点がつくりバーススコアが含まれているため、該当する項目について「8 - 回答値」の変換を行い、適切に補正する。これにより、すべてのスコアを統一した上で、全 20 項目の合計スコアを算出し、被験者の共感力を評価した。スコアの範囲は 20～140 であり、スコアが高いほど共感力が高いことを示す。

システム使用前後の JES スコアの変化を分析するため、t 検定を用いた。この解析により、システム使用前後で共感力の変化が統計的に有意であるかを検討した。

以上の分析を通じて、本システムの使用が被験者の共感力にどのような影響を与えるのかを評価し、漢方医学的問診における共感力の向上に寄与するかを検討する。

§ 5.2 実験結果と考察

本節では、5.1 節で述べた数値実験について、それぞれの結果とそれに対する考察を示す。

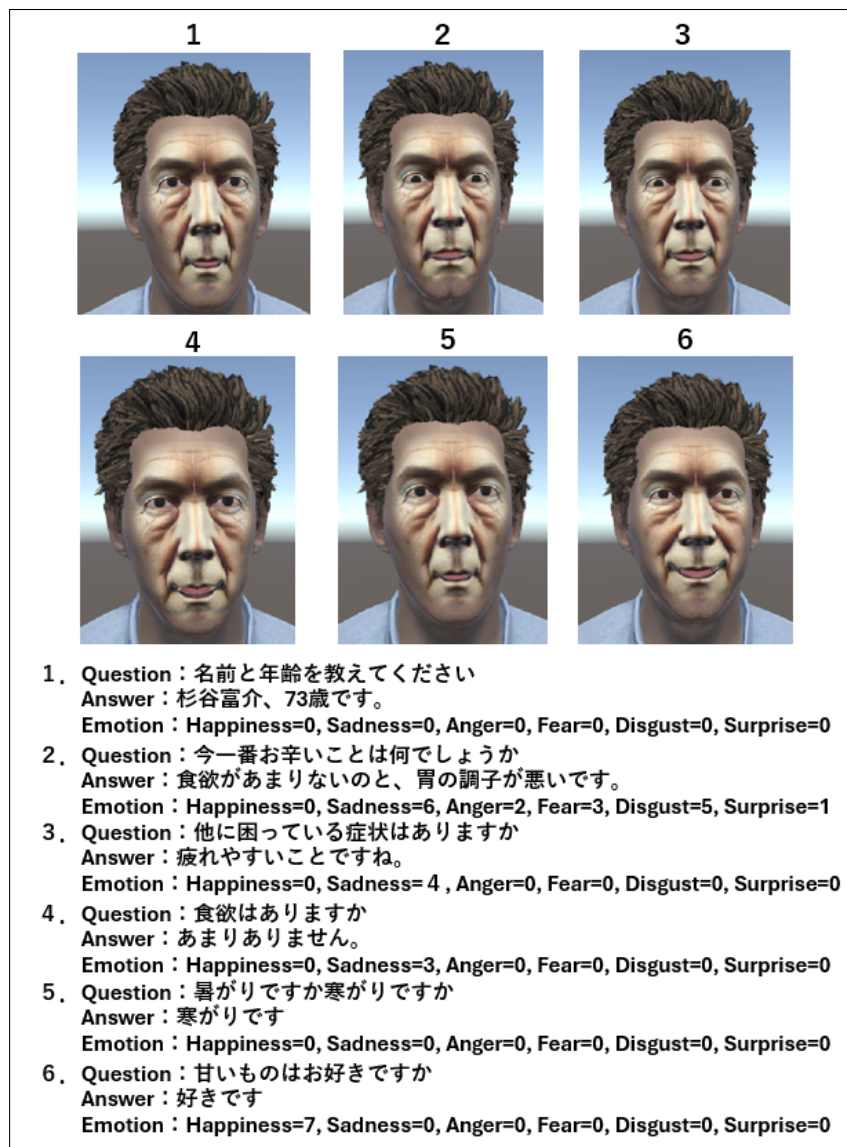


図 5.1: 出力結果

本システムの出力結果

本システムを用いて問診を行った時に出力された結果は、図5.1のようになった。最初の質問では、感情の変化は見られず、表情もニュートラルな状態である。次の質問では、「今一番辛いことは何か」という質問に対し、「食欲があまりないのと、胃の調子が悪い」と答えただけには、SadnessやDisgustが高く推定され、それに応じて表情もやや陰しく、落ち込んだような印象となっている。次に、「他に困っている症状はあるか」という質問では、「疲れやすいことですね」と回答し、Sadnessが推定されているが、数値が4であるため、表情は比較的穏やかである。また、「食欲はありますか」という質問では「あまりありません」と答え、Sadnessがやや表れているが、大きな感情変化は見られなかった。さらに、「暑がりか寒がりか」と尋ねられた際には「寒がりです」と回答したが感情スコアはすべて0であった。最後に、「甘いものは好きですか」と尋ねた際には、「好きです」と答え、Happinessが

表 5.3: t 検定の結果

	変数 1	変数 2
平均	85.3333	105.667
分散	336.667	64.6667
観測数	6	6
ピアソン相関	0.12018	
仮説平均との差異	0	
自由度	5	
t	-2.6039	
P(T<=t) 片側	0.02401	
t 境界値 片側	2.01505	
P(T<=t) 両側	0.04802	
t 境界値 両側	2.57058	

表 5.2: JES の結果

	使用前	使用后
被験者 1	85	101
被験者 2	97	107
被験者 3	113	114
被験者 4	59	114
被験者 5	81	93
被験者 6	77	105

最も高く推定された。その結果、バーチャル患者の表情も笑顔になった。

本システムの出力結果に対する考察

本システムを用いた問診の結果、学習させたシナリオに基づいた適切な回答が生成され、漢方医学的問診が行えることがわかった。さらに、本システムでは感情推定を行い、その推定値に基づいて適切な表情を選択する機能を実装しているが、実際に出力された表情は、感情値と整合性が取れており、実際の人間の感情表現に近いものとなっていたと考えられる。例えば、Sadness や Disgust のスコアが高い場合には険しい表情が生成され、Happiness が高い場合には笑顔が表現されるなど、感情に応じた表情変化が適切に反映されていた。この結果から、バーチャル患者の感情を考慮し、より共感的なシステムが構築できたと考えられる。

JES のアンケート結果

JES のアンケート結果は、表 5.2 のようになった。まず、被験者 1 は、システム使用前の JES スコアが 85 で、使用後は 101 となった。被験者 2 も同様にスコアが向上し、使用前は 97、使用後は 107 であった。被験者 3 の場合、使用前のスコアが 113 であり、使用後は 114 とわずかに上昇した。一方で、被験者 4 は使用前のスコアが 59 と比較的低かったが、使用後には 114 まで大きく改善した。被験者 5 も使用前は 81 だったが、使用後は 93 へと上昇した。さらに、被験者 6 は使用前のスコアが 77 で、使用後には 105 に向上した。

JES のアンケート結果の考察

本研究では、システム使用前後の JES スコアの変化に対して、t 検定を実施し優意水準を 0.05 に設定し、統計的有意性を評価した。表 5.3 がその結果となる。結果を確認すると、システム使用後の JES スコアの分布は使用前と比較して上昇傾向にあることが見受けられる。

特に、スコアの中央値が上昇しており、一部の被験者では顕著な改善が確認された。また、 p 値が有意水準 0.05 以下であったため、システム使用前後で統計的に有意な差があると考えられる。本結果から、バーチャル患者との対話が一定の被験者において共感力向上に寄与した可能性が示唆されるが、さらなる検証が必要であると考えられる。今後の課題として、サンプルサイズを増やし、より多様な被験者を対象とした実験を行うことで、システムの有効性をより詳細に検証する必要がある。また、被験者の事前の共感度の違いなどが結果にどのような影響を与えたのかを分析することも、今後の研究の方向性として重要であるとする。

おわりに

本研究では、生成 AI を用い、人間と同等の会話および会話が可能なバーチャル患者を構築するために、漢方専門医が実際の事例を基に作成した問診シナリオを活用し、問診でのスムーズな質問と回答を実現した。さらに、生成された回答から感情を推定し、それに応じてバーチャル患者が感情を表現できるようにすることで、実際の患者との問診に近い学習環境を提供することができた。本研究の直接的な目的である医学生の学習効果の向上を明確に実証するには至らなかったが、学習効果を検証するための基盤となる環境は構築できた。

今後は、検証実験を実施する予定であり、さらにいくつかの課題が残されている。例えば、バーチャル患者の応答速度の向上や、症状および症例のバリエーションの拡充が求められる。また、多様なシナリオに対応するためには、性別や年齢層の異なるバーチャル患者の開発が必要不可欠である。さらに、バーチャル患者の感情表現をより明確にすることも重要な課題である。これらの課題に取り組むことで、医学生は、多様な背景や特徴、医学的情報を持つ患者に対する漢方医学的問診の体験を通じて、実際の問診で活用できるスキルを習得できると考えられる。

また、医学生の共感の低下が指摘され、その育成方法が国際的にも模索される状況を踏まえ、本研究は医学生の共感を向上させる新たな教育方法として、医学教育に貢献できると考える。さらに、本研究ではバーチャル患者とのインタラクションを優先したため、現段階では1つの問診シナリオしか構築できていない。今後は、より多様な患者の属性および主訴や症状などの医学的情報を医学生が体験できるように、複数のシナリオを開発することに注力する必要がある。このようなシステムの拡張により、医学生は臨床実習で体験できない事例に対する漢方医学的問診を体験することができ、彼らの共感能力および漢方医学的病態の理解が促進されることが期待できる。

さらに、本研究で開発したバーチャル患者の実用性と教育効果を検証するため、医学生を対象とした行動実験を計画している。この実験では、バーチャル患者に問診を行う医学生と、模擬患者に問診を行う医学生の2群にランダムに割り付け、共感得点、漢方医学的問診スキル、漢方医学的病態の理解度を指標として評価する予定である。

謝辞

本研究を遂行するにあたり，多大なご指導と終始懇切丁寧なご鞭撻を賜った富山県立大学工学部情報工学部データサイエンス学科システム数理講座の António Oliveira Nzinga René 講師，奥原浩之教授に深甚な謝意を表します．また，研究に協力していただいた富山大学大学院医学薬学研究成人看護学の山田理絵准教授，富山大学学術研究部医学系藤本誠准教授，弘前大学教育推進機構の徐助教，日本漢方医学教育財団様に感謝の意を表します．最後になりましたが，多大な協力をしていただいた研究室の同輩諸氏に感謝いたします．

2025 年 2 月

高田知樹

参考文献

- [1] G.D. Shumer, S.L. Warber, S. Motohara, A. Yajima, M.A. Plegue, M.F. Bialko, T. Iida, K. Sano, M. Amenomori, T. Tsuda, & M.D. Fetters, “Complementary and alternative medicine use by visitors to rural Japanese family medicine clinics: results from the international complementary and alternative medicine survey,” *BMC Complementary and Alternative Medicine*, Vol. 14, 2014.
- [2] P.S. Bellet & M.J. Maloney, “The importance of empathy as an interviewing skill in medicine,” *JAMA*, Vol. 266, No. 13, pp. 1831–1832, 1991.
- [3] Q. Wu, Z. Jin, & P. Wang, “The Relationship Between the Physician-Patient Relationship, Physician Empathy, & Patient Trust,” *Journal of General Internal Medicine*, Vol. 37, pp. 1388–1393, 2021.
- [4] M.A. Quinn, L.M. Grant, E. Sampene, & A.B. Zelenski, “A Curriculum to Increase Empathy and Reduce Burnout,” *WMJ: Official Publication of the State Medical Society of Wisconsin*, Vol. 119, No. 4, pp. 258–262, 2020.
- [5] M. Hojat, D.Z. Louis, K. Maxwell, F.W. Markham, R.C. Wender, & J.S. Gonnella, “Patient perceptions of physician empathy, satisfaction with physician, interpersonal trust, & compliance,” *International Journal of Medical Education*, Vol. 1, pp. 83–87, 2020.
- [6] E. Larson & X. Yao, “Clinical empathy as emotional labor in the patient-physician relationship,” *JAMA*, Vol. 293, No. 9, pp. 1100–1106, 2005.
- [7] M. Flook, C. Jackson, E. Vasileiou, C.R. Simpson, M.D. Muckian, U. Agrawal, C. McCowan, Y. Jia, J.L.K. Murray, L.D. Ritchie, C. Robertson, S. Stock, J.X. Wang, M.E.J. Woolhouse, A. Sheikh, & H.R. Stagg, “Informing the public health response to COVID-19: a systematic review of risk factors for disease, severity, & mortality,” *BMC Infectious Diseases*, Vol. 21, 2021.
- [8] S. Mao, L. Guo, P. Li, K. Shen, M. Jiang, & Y. Liu, “New era of medical education: asynchronous and synchronous online teaching during and after COVID-19,” *Advances in Physiology Education*, Vol. 47, pp. 272–281, 2023.
- [9] M. Bearman, B. Cesnik, & M.J. Liddell, “Random comparison of ‘virtual patient’ models in the context of teaching clinical communication skills,” *Medical Education*, Vol. 35, pp. 824–834, 2001.
- [10] A. Foster, J. Harms, B.L. Ange, B. Rossen, B.C. Lok, S. Lind, & C.L. Palladino, “Empathic Communication in Medical Students’ Interactions with Mental Health Virtual Patient Scenarios: A Descriptive Study Using the Empathic Communication Coding System,” *Austin J. Psychiatry Behav Science*, 2014.

- [11] A. Foster, N. Chaudhary, T. Kim, J.L. Waller, J. Wong, M. Borish, A. Cordar, B.C. Lok, & P.F. Buckley, “Using Virtual Patients to Teach Empathy: A Randomized Controlled Study to Enhance Medical Students’ Empathic Communication,” *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, Vol. 11, pp. 181–189, 2016.
- [12] T.C. Guetterman, R. Sakakibara, S. Baireddy, F.W. Kron, M.W. Scerbo, J.F. Cleary, & M.D. Fetters, “Medical Students’ Experiences and Outcomes Using a Virtual Human Simulation to Improve Communication Skills: Mixed Methods Study,” *Journal of Medical Internet Research*, Vol. 21, 2019.
- [13] W.T. Gunning & U.G. Fors, “Virtual Patients for assessment of medical student ability to integrate clinical and laboratory data to develop differential diagnoses: Comparison of results of exams with/without time constraints,” *Medical Teacher*, Vol. 34, pp. 222–228, 2012.
- [14] A. Isaza-Restrepo, M.T. Gómez, G. Cifuentes, & A. Argüello, “The virtual patient as a learning tool: a mixed quantitative qualitative study,” *BMC Medical Education*, Vol. 18, No. 1, p. 297, 2018.
- [15] NEC ソリューションイノベータ, “生成 AI とは？従来の AI との違いや企業活用のメリットを解説”, 閲覧日 2025-02-01
https://www.necsolutioninnovators.co.jp/sp/contents/column/20240426_generativeai-ai.html.
- [16] R. Yamada, K. Xu, M. Fujimoto, & T. Kido, “Gaze of Kampo specialists during simulated dynamic abdominal examinations: A cross-sectional study using a wearable eye tracker,” *Traditional & Kampo Medicine*, Vol. XX, pp. 1–10, 2024.
- [17] Reinforz Insight, “AI と教育の融合：個別最適化学習で実現する未来の学習環境”, 閲覧日 2025-02-01
<https://reinforz.co.jp/bizmedia/50729/>.
- [18] AI 総研, “【事例 5 選】生成 AI を教育現場に導入する 7 大メリットとは?”, 閲覧日 2025-02-01
https://metaversesouken.com/ai/generative_ai/education/.
- [19] ベネッセ教育情報サイト, “自由研究のお困りは「自由研究おたすけ AI β 版」で解決!”, 閲覧日 2025-02-01
<https://benesse.jp/kyouiku/202308/20230815-2.html>.
- [20] 株式会社学研ホールディングス, “学研オリジナル学習システム (GDLS) で ChatGPT を活用し、生徒の学習効果を最大化する個別アドバイスを提供開始”, 閲覧日 2025-02-01
https://www.gakken.co.jp/ja/news_item/2023/202320230712102949.html.

- [21] 文部科学省, “ 初等中等教育段階における生成 AI の利用に関する暫定的なガイドライン ”, 閲覧日 2025-02-01
https://www.mext.go.jp/content/20230710-mxt_shuukyo02-000030823_003.pdf.
- [22] AI Market, “ 医療分野での生成 AI (ジェネレーティブ AI) 活用法や事例、メリット、注意点を解説！医療品質の向上に生成 AI はどう使われる？”, 閲覧日 2025-02-01
<https://ai-market.jp/industry/generative-ai-medical/>.
- [23] Ubie 株式会社のプレスリリース, “ 恵寿総合病院と Ubie、生成 AI を活用した「医師の働き方改革」の実証実験を実施 ”, 閲覧日 2025-02-01
<https://prtimes.jp/main/html/rd/p/000000066.000048083.html>.
- [24] 株式会社 HOKUTO のプレスリリース, “ 4 月 3 日、医師向け臨床支援アプリ「HOKUTO」に導入開始 AI 技術 OpenAI GPT-4 を活用した新機能 ～患者への説明内容の考案を支援、キーワードから最新のおすすめ研究論文を抽出～”, 閲覧日 2025-02-01
<https://prtimes.jp/main/html/rd/p/000000012.000030964.html>.
- [25] M. Karabacak, B.B. Ozkara, K. Margetis, M. Wintermark, & S. Bisdas, “ The Advent of Generative Language Models in Medical Education,” JMIR Medical Education, Vol. 9, p. e48163, 2023.
- [26] C-MECJP, “ 臨床実習とは？実習の種類や目的を確認しよう ”, 閲覧日 2025-02-01
<https://c-mec.jp/notice/docs0006htnt/>.
- [27] 元濱 奈穂子, “ 医学教育における模擬患者との「協働」の実態 ”, 教育社会学研究 Vol. 106, pp. 93-114, 2021.
- [28] M. Benfatah, A. Marfak, E. Saad, A. Hilali, C. Nejari, & I. Youlyouz-Marfak, “ Assessing the efficacy of ChatGPT as a virtual patient in nursing simulation training: A study on nursing students’ experience,” Teaching and Learning in Nursing, Vol. 19, No. 3, pp. 486–493, 2024.
- [29] I. Gonullu, A. Bayazit, & S. Erden, “ Exploring medical students’ perceptions of individual and group-based clinical reasoning with virtual patients: a qualitative study,” BMC Medical Education, Vol. 24, No. 1, p. 189, 2024.
- [30] SPJ, “ 対話システムを構成する 2 つの仕組みと、フレームワークとは? ”, 閲覧日 2025-02-01
<https://spjai.com/interactive-system/>.
- [31] 小林 峻也, 萩原 将文, “ ユーザの嗜好や人間関係を考慮する非タスク指向型対話システム ”, 人工知能学会論文誌, Vol. 31, No. 1, pp. 1–10, 2016.
- [32] NTT, “ 機械が会話のパートナーになる日 ——大規模深層学習で拓く雑談対話システムの新境地 ”, 閲覧日 2025-02-01
https://www.rd.ntt/research/JN202107_14467.html.

- [33] NTT 東日本, “AI 音声認識とは？仕組み・活用方法を分かりやすく解説”, 閲覧日 2025-02-01
<https://business.ntt-east.co.jp/content/cloudsolution/column-253.html#section-03>.
- [34] AmiVoice Techblog, “音声認識の仕組みをざっくり解説!”, 閲覧日 2025-02-01
<https://acp.amivoice.com/blog/2023-03-06-094430/>.
- [35] A. Radford, J.W. Kim, T. Xu, G. Brockman, C. McLeavey, & I. Sutskever, “Robust Speech Recognition via Large-Scale Weak Supervision,” arXiv preprint, 2022.
- [36] OTONARI, “音声合成の仕組みをわかりやすく解説!”, 閲覧日 2025-02-01
<https://www.soft-voicesynthesis.com/knowledge/structure.html>.
- [37] AI 総合研究所, “LSTM とは？その仕組みや使用例をわかりやすく解説!”, 閲覧日 2025-02-01
<https://www.ai-souken.com/article/lstm-overview>.
- [38] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, L. Kaiser, & I. Polosukhin, “Attention Is All You Need,” arXiv preprint, 2017.
- [39] ainow, “ファインチューニングとは？手順やメリット、転移学習の違いを解説”, 閲覧日 2025-02-01
<https://ainow.jp/fine-tuning/>.
- [40] A. Radford, K. Narasimhan, T. Salimans, & I. Sutskever, “Improving Language Understanding by Generative Pre-Training,” OpenAI, 2018.
- [41] AIsmiley, “プロンプトエンジニアリングとは？ChatGPT で代表的な 12 個のプロンプトを実演!”, 閲覧日 2025-02-01
https://aismiley.co.jp/ai_news/what-is-prompt-engineering/.
- [42] M. García Corretjer, D. Miralles, & R. Ros, “A Theoretical Approach for a Novel Model to Realizing Empathy,” arXiv preprint, 2020.
- [43] 梅田 聡, “共感の理論と脳内メカニズム”, 高次脳帰納研究 Vol. 38, No. 2, pp. 133-138, 2018.
- [44] 長峯 正典, 佐野 信也, 重村 淳, 吉野 相英, 清水 邦夫, “医療従事者に求められる共感性に関する考察”, 日本総合病院精神医学会誌, Vol. 31, No. 2, pp. 147-153, 2019.
- [45] P. Menezes, S. S. Guraya, and S. Y. Guraya, “A Systematic Review of Educational Interventions and Their Impact on Empathy and Compassion of Undergraduate Medical Students,” Frontiers in Medicine, vol. 8, 2021.
- [46] Mark H. Davis, “Measuring individual differences in empathy: Evidence for a multi-dimensional approach,” Journal of Personality and Social Psychology, Vol. 44, No. 1, pp. 113–126, 1983.

- [47] Robert Hogan, "Development of an empathy scale," *Journal of Consulting and Clinical Psychology*, Vol. 33, No. 3, pp. 307–316, 1969.
- [48] A. Mehrabian and N. Epstein, "A Measure of Emotional Empathy," *Journal of Personality*, Vol. 40, No. 4, pp. 525–543, 1972.
- [49] M. Hojat, J.S. Gonnella, T.J. Nasca, S. Mangione, J. Vergare, & M. Magee, "The Jefferson Scale of Physician Empathy: Development and Preliminary Psychometric Data," *Educational and Psychological Measurement*, Vol. 61, No. 2, pp. 349–365, 2001.
- [50] Google Cloud, "Speech-to-Text AI: 音声認識と音声文字変換", 閲覧日 2025-02-01
<https://cloud.google.com/speech-to-text?hl=ja>.
- [51] VOICEVOX, "無料で使える中品質なテキスト読み上げソフトウェア", 閲覧日 2025-02-01
<https://voicevox.hiroshiba.jp/>.
- [52] 深尾遼平, 野瀬裕太, 三角昌詩, 青木秀郎, 小島亮太, 岩田宜久, 松本和久, "東洋医学的診察の西洋医学的診察との整合性と現代医療における臨床的意義", *日本東洋医学研究会誌*, vol. 2, pp. 21–28, 2016.
- [53] OpenAI, "ChatGPT", 閲覧日 2025-02-01
<https://openai.com/ja-JP/chatgpt/overview/>.
- [54] OpenAI, "Fine-tuning", 閲覧日 2025-02-01
<https://platform.openai.com/docs/guides/fine-tuning>.
- [55] XR CLOUD, "メタバースと VR の違いとは? ビジネス活用の可能性も解説", 閲覧日 2025-02-01
<https://xrcloud.jp/blog/articles/business/6652/>.
- [56] Reallusion, "Character Creator", 閲覧日 2025-02-01
<https://www.reallusion.com/jp/character-creator/default.html>.
- [57] P. Ekman & W.V. Friesen, "Constants across cultures in the face and emotion," *Journal of Personality and Social Psychology*, Vol. 17, No. 2, pp. 124–129, 1971.
- [58] J. Broekens, B. Hilpert, S. Verberne, K. Baraka, P. Gebhard, and A. Plaat, "Fine-grained Affective Processing Capabilities Emerging from Large Language Models," *arXiv preprint arXiv:2309.01664*, 2023.
- [59] N. Wake, A. Kanehira, K. Sasabuchi, J. Takamatsu, & K. Ikeuchi, "Bias in Emotion Recognition with ChatGPT," *arXiv preprint, arXiv:2310.11753*, 2023.
- [60] P. Ekman, "An Argument for Basic Emotions," *Cognition and Emotion*, Vol. 6, No. 3-4, pp. 169–200, 1992.

- [61] 片岡 仁美, 「共感と医療について (エンパシースケールを中心に)」, 日本内科学会雑誌, 第 101 巻, 第 7 号, pp. 2103–2107, 2012.
- [62] Unity, “リアルタイム開発プラットフォーム — 2D/3D、VR/AR エンジン”, 閲覧日 2025-02-03
<https://unity.com/ja>.
- [63] Uezo, “ChatdollKit: A framework for developing interactive chatbot applications in Unity ”, 閲覧日 2025-02-03
<https://github.com/uezo/ChatdollKit>.