大規模言語モデルは日本語・中国語の状態パーフェクトを理解できるか?

盧 捷 [1] 金 杜 [1] 柴田 行輝 [1] 土井 智暉 [1] 染谷 大河 [1] 谷中 瞳 [1]
[1] 東京大学
lu-jie20010825@g.ecc.u-tokyo.ac.jp

概要

近年，大規模言語モデル (Large Language Model, LLM) は様々な言語理解タスクで高精度を達成しつつある．しかし，LLM が文中に現れる事象間の時間関係をどの程度理解できているのかについては，十分に研究が進められていない．本研究では，「パーフェクト相」の一種である「状態パーフェクト」に注目して自然言語推論（Natural Language Inference, NLI）データセットを構築し，LLM の日本語・中国語の状態パーフェクトが表す複数の時点の関係性を正確に把握しているかについて分析を行った．実験の結果，日本語・中国語の状態パーフェクトの表す意味について，LLM の判断は一貫しておらず，人間の判断と一致していない傾向を明らかにした．

1 はじめに

近年，大規模言語モデル (Large Language Model, LLM) はさまざまな自然言語理解タスクのベンチマークで高精度を達成しつつある．しかし，言語モデルが自然言語文に現れる時間関係をどの程度認識できているのかについては自明でない．テンス (時制)・アスペクト (相) は時間を捉えるための中核的な言語現象であり，LLM がどの程度テンス・アスペクトを理解しているかという問いに対して関心が寄せられている．

近年の研究として，[1] と [2] はそれぞれ言語モデルの日本語と英語のテンス・アスペクトの理解度を分析している．しかし，これまでの日本語・中国語を対象とした研究の中に，「**パーフェクト相 (完了相)**」を対象としたものはまだない．「パーフェクト相」は，アスペクトの一種であり，文中の動作または出来事の効力が，時間軸上に基準になる時点 (参照時) まで存続しているという意味を表す．英語の完了形は，典型的なパーフェクト相である．日本語と中国語において，英語の have のようなパーフェクト相専用の文法的標識がないため，パーフェクト的意味は基本「テイル」「テイタ」[3] と「了 (中国語)」[4] が担っている．

(1) a. 花子が先週帰宅したとき，太郎はすでに**死んでいた**.
b. 花子上周回到家的 时候，太郎已 经死**了**.
c. When Hanako returned home last week, Taro **had already died**.

LLM の言語理解能力を評価する主要なタスクの１つとして，自然言語推論（Natural Language Inference，NLI）がある．NLI とは，与えられた前提文（Premise）が真であるとき，仮説文（Hypothesis）もまた真である場合は含意（Entailment）を，必ずしも真とはいえない場合は非含意（Non-Entailment）を判定するタスクである．

もしモデルが日本語のパーフェクト相の表す時間的意味を理解できているならば，以下の 2 例に対して正しいラベルを出力するはずである．

(2) P: 太郎が３日前に退職した.
H: 太郎が退職している．（**Entailment**）

(3) P: 花子が来週帰国する時，太郎は三日前に退職している.
H: 太郎が退職している．（**Non-Entailment**）

本研究は，日本語・中国語を対象言語として，NLI タスクを通じて，**LLM が状態パーフェクト (2.1節参照) が表している複数の時点の関係性を正確に把握しているかについて分析する**.

本研究の貢献は主に以下の二点である．第一に，本研究は，日本語および中国語の状態パーフェクトに注目した初めての NLI データセットを構築した．第二に，前述したデータセットを用いて `GPT-4` および `Claude-3.5` の評価を行い，日中両言語の状態パーフェクトについて，LLM の判断は一貫しておらず，人間の判断と一致していない傾向を明らかにした．

2 背景

2.1 パーフェクト相と文中の時間関係

本節では，パーフェクト相とそれに関わる文中の時間関係について説明する．

まず, 動詞の種類によって，パーフェクト相は変化動詞（例: 死ぬ）による「**状態パーフェクト**」と動作動詞（例:「走る」）による「**動作パーフェクト**」の２種類に分けられ，それぞれ異なる時間関係を表している．本研究は，含意関係がなるべく一意に定まるよう，結果状態に終了限界がない変化動詞変化動詞による「**状態パーフェクト**」のみを対象とする．

文中の時間関係は，文が発された時点である発話時 (speech time)，時間軸上の基準になる参照時 (reference time)，変化が生じた出来事時 (event time) という３つの時点を使用して記述することができる [5]. 特に，状態パーフェクトにおける三つの時点の関係について，以下のことが知られている [6][3][7].

(a) 発話時は常に「現在」に置かれるが，参照時は文脈の指定によって過去・現在・未来のいずれにもなりうる. (無指定の場合は現在)

(b) 出来事時は常に参照時より先行している．

例えば，(4a) においては文脈上の指定がないため，参照時は発話時と一致し「現在」に設定されるのに対して，(4b)(4c) においては，参照時はそれぞれ過去，未来に設定され，いずれも時間軸において出来事時は参照時に先行している．このように，状態パーフェクトにおける全体の時間関係は参照時によって決まるので，状態パーフェクトは参照時が過去・現在・未来のいずれであるかによって，３種類に分けられる．

(4) a. 鈴木は卒業している．
出来事時→参照時=発話時

b. 山田が去年家を建てた時，鈴木はすでに卒業していた．
出来事時→参照時→発話時

c. 花子が来年家を建てる時，鈴木はとっくに卒業している．
発話時→出来事時→参照時 または
出来事時→発話時→参照時

また，(5) のように，参照時と出来事時の時間差が文中に**時間副詞**として明示される文も存在する．

(5) 山田が去年家を建てた時，鈴木は既に**３ヶ月前に**卒業していた．
出来事時→３ヶ月経過→参照時→発話時

以上より，状態パーフェクトの文中の時間関係を扱うには「**参照時 (過去/現在/未来)**」及び「**時間副詞 (有/無)**」という二つの要素が重要であることがわかる．ここまで，日本語を中心に議論してきたが，これらの特性は中国語にも共通している [4].

2.2 時間推論データセット

日本語に関する LLM の時間推論能力を分析する先行研究として，[1] はアスペクト関連の問題も含めた日本語の時間推論データセットを構築し，LLM の評価を行なっている．この研究では，`GPT-3.5` と `GPT-4` がアスペクト関連の問題において精度が低いと報告しているが，パーフェクト相に関する体系的な議論は行われていない．一方，中国語の NLI データセットとしては，OCNLI[8] の一部に時間推論の問題も含まれているが，LLM の時間推論能力については十分に分析が進められていない．

本研究の対象言語でないが，英語圏ではパーフェクト相も含めて英語の複数のテンス・アスペクト現象を扱った NLI データセットがテンプレートベースで構築されている [2]．本研究では，この研究に着想を得てテンプレートベースで日本語・中国語の状態パーフェクトに関する NLI データセットを構築し，現在の LLM がどの程度体系的に状態パーフェクトを扱っているのか分析を行う．

3 データセットの構築

本研究では2.1節を踏まえて，「**参照時が過去/現在/未来**」及び「**時間副詞 (有/無)**」の二つの要素を考慮して, 各言語（日本語・中国語）に対して文テンプレートを 6 種類人手で作成した．これらのテンプレートを NLI タスクの前提・仮説に割り当てることで，30 件の推論テンプレートを作成した．さらに，推論テンプレートに割り当てる語彙の組み合わせを各言語につき 45 件作成した.[1)] 最終的に，各言語ごとに 1350 件（= 30 × 45）の問題が構築された．

文テンプレートと対応する例文は，表 1(日本語) と付録の表7（中国語）で示す. 問題と正解ラベルの例文は付録の表6に示す.

時間推論に関する問題を作成する際，文脈やテン

1) 言語間比較のため，推論テンプレートと割り当てる語彙は日中両言語の意味が対応するものにしている．

表 1　時間関係と対応する例文 (日本語) の一覧. 時間関係において，(副詞) の有無は時間副詞の有無を示す. この例では，テンプレートに割り当てる項目として，[event-past]:「田中が先月引っ越した」，[event-future]:「佐藤が来月転職する」，[np]:「山本」，[time]:「一週間」，[v-teiru]:「合格している」を使用している.

時間関係	文テンプレート	例文（日本語）
過去 (副詞)	[event-past] とき、[np] は [time] 前にすでに [v-teita]	田中が先月引っ越したとき、山本は一週間前にすでに大学に合格していた
過去	[event-past] とき、[np] はすでに [v-teita]	田中が先月引っ越したとき、山本はすでに大学に合格していた
現在 (副詞)	[np] は [time] 前に [v-ta]	山本は一週間前に大学に合格した
現在	[np] は [v-teiru]	山本は大学に合格している
未来 (副詞)	[event-future] とき、[np] は [time] 前に [v-teiru]	佐藤が来月転職するとき、山本は一週間前に大学に合格している
未来	[event-future] とき、[np] はとっくに [v-teiru]	佐藤が来月転職するとき、山本はとっくに大学に合格している

プレートに割り当てる語彙などによって，文が不自然になる場合や，正解の含意関係ラベルが一意に決まらない場合がある. これらの問題点に対処するために，本研究では，問題を作成する際に以下の 2 点に留意している.

(i) 本研究の問題の前提・仮説に使用する各言語の文の自然さは，各言語の母語話者が人手でチェックし，不自然なものを修正した.

(ii) 正解ラベルの妥当性を検証するために，推論テンプレートをすべて含んだ 30 問の問題をアンケートにして，各言語の複数名[2] の母語話者に解くように依頼した. 多数決のもとで，アンケートの全問において，母語話者の判断が正解の含意関係ラベルと一致していることを確認した.

4　実験

各モデルに説明用の指示プロンプトと前提・仮説を入力として与え，正答率で評価を行った.

4.1　実験設定

モデル　多言語の LLM として `GPT-4`(gpt-4-0613) および `Claude-3.5`（claude-3-5-sonnet-20241022）を評価対象とした.

指示プロンプト　指示プロンプトは，先行研究 [1] の Baseline プロンプトを参考にして作成した. 各言語の指示プロンプトは付録の表8に示す. 全ての実験は，zero-shot で行った.

4.2　結果と考察

すべての問題の平均正答率 (%) を集計したところ，日本語，中国語に対して `Claude-3.5` はそれぞれ 88.4%と 91.5%，`GPT-4` は 78.4%，80.6%の正答率だった. この結果について，「時間副詞の有無」と「参照時」を軸に，各推論問題におけるモデルの予測の傾向を分析する. 以降では便宜上，前提と仮説の時間関係をもとに問題を記述する. 例えば，(P: 過去 (副詞)，H: 未来) は，前提と仮説がそれぞれ「過去 (副詞)」と「未来」の例文の問題を指す.

2)　日本語母語話者 3 名，中国語母語話者 7 名に依頼した.

表 2　参照時が同じで，時間副詞の有無だけが異なる問題の結果. 結果は「日本語の正答率 (%)：中国語の正答率 (%)」で示す. 両モデルとも高い正答率を示した.

前提	仮説	正解	GPT-4	Claude
過去 (副詞)	過去	N	100:100	100:100
過去	過去 (副詞)	E	95.6:100	100:100
現在 (副詞)	現在	N	100:100	100:100
現在	現在 (副詞)	E	100:100	100:100
未来 (副詞)	未来	N	97.8:100	100:100
未来	未来 (副詞)	N	86.7:97.8	100:100

表 3　(i) (P: 過去/未来,H: 現在) の実験結果. 結果は「日本語の正答率 (%)：中国語の正答率 (%)」で示す.「正解」は元の問題の正答率を，「主節」は前提を元の問題の前提の主節のみに置き換えた問題の正答率である. *E* と *N* はそれぞれ **E**ntailment と **N**on-Entailment を意味する.

前提	仮説	正解	主節	GPT-4	Claude-3.5
過去 (副詞)	現在 (副詞)	N	E	**4.4:0.0**	**33.3:77.8**
	現在	E	E	100:100	93.3:100
過去	現在 (副詞)	N	N	100:100	97.8:100
	現在	E	E	97.8:100	75.6:95.6
未来 (副詞)	現在 (副詞)	N	E	**4.4:6.7**	**75.6:91.1**
	現在	N	E	**0.0:0.0**	**73.3:53.3**
未来	現在 (副詞)	N	N	100:100	100:100
	現在	N	E	**2.2:0.0**	**95.6:86.7**

まず，前提と仮説の参照時が同じ，かつ時間副詞の有無だけが異なる問題においては，両モデルとも 100%に近い正答率を示した (表 2参照). この結果から，`Claude-3.5`，`GPT-4` が参照時が同じ場合の時間副詞の有無によるによる意味の違いをある程度認識できていることが示唆された.

次に, 前提と仮説の参照時が異なる場合の結果から，以下の 3 つの傾向が示唆された.

(i) (P: 過去/未来,H: 現在)　(P: 過去/未来,H: 現在) の問題において，パーフェクト相を理解できる LLM なら，主節の出来事時が常に従属節の参照時に先行すると認識するはずだが，`GPT-4` は，前提の主節と従属節の時間関係をうまく認識できず，前提の主節の情報のみに従ってラベルを予測している可能性

表 4 (ii) (P: 過去 (未来),H: 未来 (過去)) の実験結果. 結果は「日本語の正答率 (%)：中国語の正答率 (%)」で示す. Claude-3.5 はこれらの問題で一貫して Non-Entailment を予測する傾向がある.

前提	仮説	正解	GPT-4	Claude-3.5
過去 (副詞)	未来 (副詞)	N	97.8:100	100:100
	未来	E	**95.6:91.1**	**62.2:53.3**
過去	未来 (副詞)	N	97.8:88.9	100:100
	未来	E	**86.7:71.1**	**55.6:35.6**
未来 (副詞)	過去 (副詞)	N	88.9:100	100:100
	過去	N	80.0:86.7	100:100
未来	過去 (副詞)	N	100:100	97.8:100
	過去	N	97.8:100	100:100

表 5 (iii) (P: 現在,H: 過去/未来) の実験結果. 結果は「日本語の正答率 (%)：中国語の正答率 (%)」で示す.

前提	仮説	正解	GPT-4	Claude
現在 (副詞)	過去 (副詞)	N	86.7:95.6	97.8:100
	過去	N	93.3:77.8	97.8:100
現在	過去 (副詞)	N	68.9:95.6	93.3:100
	過去	N	82.2:55.6	93.3:97.8
現在 (副詞)	未来 (副詞)	N	**33.3:82.2**	**22.2:97.8**
	未来	E	100:100	100:86.7
現在	未来 (副詞)	N	**55.6:68.9**	**84.4:100**
	未来	E	100:100	100:68.9

がある (表3参照). 具体的に，正解ラベルが元の問題の主節のみを前提とした場合の正解ラベルと一致する場合は，すべてのモデルにおいて非常に高い正答率を示したのに対して，一致していない場合は，GPT-4 はほぼ不正解のラベルを予測し，Claude-3.5 も一部の問題において正答率がチャンスレートまで低下している. 例えば，(6) の (現在, 未来) の問題において，正解は Non-Entailment であるが，主節のみを前提とした場合 Entailment になる.

(6) P: 田中が来月引っ越すとき，**山本はとっくに大学に合格している**. (未来)
H: 山本は大学を合格している. (現在)
正解:Non-Entailment
前提の主節だけで判断: Entailment

(ii) (P: 過去,H: 未来) & (P: 未来,H: 過去) これらの問題において，Claude-3.5 は Non-Entailment を予測する傾向がある. 正解が Non-Entailment の問題において，両モデルとも高い正答率を示したが，Claude-3.5 は正解が Entailment である問題において, 正答率がチャンスレート付近で変動しており，GPT-4 も (過去, 未来) の問題で正答率が 70%まで低下している.

(iii) (P: 現在，H: 過去/未来) (P: 現在，H: 過去/未来) の問題において，GPT-4 と Claude-3.5 は (現在 (副詞), 未来 (副詞)) 及び (現在, 未来 (副詞)) の 2 問で正答率が低下する傾向がある (表5参照). とくに，(現在 (副詞), 未来 (副詞)) において，日本語と中国語の正答率の差は両モデルとも約 50%あった. 両モデルともこれほどの正答率の差が出たのはこの一問だけであるため，これはモデルの性能によるものでなく，言語の差に由来すると考える. 日本語での正答率が低い原因として，LLM が日本語の仮説に現れる未来 (副詞) の時間副詞を従属節の参照時を修飾する表現と誤って解釈したことが考えられる. このとき，仮説の参照時は従属節の示す時点でなく，従属節の時点から時間副詞の示す時間よりも前の時点となり，例えば (7) の表す意味を (7a) ではなく (7b) と認識してしまう. 仮説をこのように解釈すると，現在 (副詞) の前提に対して不正解の Entailment が予測され，正答率が低下したと考えられる.

(7) 山田が来週本を出版するとき，鈴木は三日前に入学試験に合格**している**.
a. 山田が来週本を出版するときの三日前に，鈴木は入学試験に合格**する**.
b. 山田が来週本を出版するときの三日前に，鈴木はとっくに入学試験に合格**している**.
出来事時 (合格) → 参照時 (本を出版するときの三日前)

なお，一つのモデルでしか正答率の差がみられなかった問題については，言語の性質でなくモデルの性能による可能性があるので，本研究では分析対象としなかった．この分析は今後の研究に譲りたい.

5 おわりに

本研究は，日本語と中国語における状態パーフェクト相に着目し，NLI データセットを構築して GPT-4 および Claude-3.5 の性能評価を実施した. 実験の結果から，参照時が同一の場合は，LLM は時間副詞の有無による意味の差異を適切に認識できている傾向がみられたが，参照時が異なる文においては人間の判断と乖離があることが確認され，LLM が状態パーフェクトの意味を一貫して認識することができていない可能性が示された. また，日本語と中国語の正答率に大きな差が見られた問題について，両言語の特徴に基づいて分析を行った. 今後，動作パーフェクトも含めたデータセットを構築するなど，LLM のさらなる評価と分析を進める.

謝辞

本研究の執筆と分析にあたり，小西いずみ先生，松岡大樹さんから有益な助言をいただきました．この場を借りてお礼申し上げます。また、推論問題の作成に際し、アンケートにご協力いただきました皆様にも深く御礼申し上げます。本研究は JST さきがけ JPMJPR21C8 の支援を受けたものである．

参考文献

[1] 杉本智紀, 尾上康雅, 谷中瞳. 「アスペクトを考慮した日本語時間推論データセットの構築」. ジャーナル フリー, Vol. 31, No. 2, pp. 637–679, 2024.

[2] Thomas Kober, Sander Bijl de Vroe, and Mark Steedman. Temporal and aspectual entailment. In **Proceedings of the 13th International Conference on Computational Semantics - Long Papers**, pp. 103–119, Gothenburg, Sweden, May 2019. Association for Computational Linguistics.

[3] 工藤真由美. 『アスペクト・テンス体系とテクスト：現代日本語の時間の表現』. 日本語研究叢書, 第 2 期第 7 巻. ひつじ書房, 11 1995.

[4] 望月圭子. 「中国語のパーフェクト相」. 『東京外国語大学論集』, Vol. 55, pp. 55–71, 1997.

[5] Hans Reichenbach. **Elements of Symbolic Logic**. Macmillan & Co., New York, 1947.

[6] 須田義治. 「シテイル形の表すパーフェクト的な意味」. 『大東文化大学紀要. 人文科学』, Vol. 61, pp. 173–191, 2 2023. 著者版フラグ: publisher.

[7] 金水敏, 工藤真由美, 沼田善子. 『時・否定と取り立て』, 第 2 巻. 岩波書店, 東京, 2000.

[8] Hai Hu, Kyle Richardson, Liang Xu, Lu Li, Sandra Kuebler, and Larry Moss. Ocnli: Original chinese natural language inference. In **Findings of EMNLP**, 2020.

A 付録

表 6 前提, 仮説の例文と含意ラベルの一覧.

前提	仮説	例文	ラベル
過去 (副詞)		山田が去年家を建て終わったとき，鈴木は三ヶ月前にすでに大学を卒業していた	
	過去	山田が去年家を建て終わったとき，鈴木はすでに大学を卒業していた	Entailment
	現在 (副詞)	鈴木は三ヶ月前に大学を卒業した	Non-Entailment
	現在	鈴木は大学を卒業している	Entailment
	未来 (副詞)	来年佐藤の就職先が決まるとき，鈴木は三ヶ月前に大学を卒業している	Non-Entailment
	未来	来年佐藤の就職先が決まるとき，鈴木はとっくに大学を卒業している	Entailment
過去		山田が去年家を建て終わったとき，鈴木はすでに大学を卒業していた	
	過去 (副詞)	山田が去年家を建て終わったとき，鈴木は三ヶ月前にすでに大学を卒業していた	Non-Entailment
	現在 (副詞)	鈴木は三ヶ月前に大学を卒業した	Non-Entailment
	現在	鈴木は大学を卒業している	Entailment
	未来 (副詞)	来年佐藤の就職先が決まるとき，鈴木は三ヶ月前に大学を卒業している	Non-Entailment
	未来	来年佐藤の就職先が決まるとき，鈴木はとっくに大学を卒業している	Entailment
現在 (副詞)		鈴木は三ヶ月前に大学を卒業した	
	過去 (副詞)	山田が去年家を建て終わったとき，鈴木は三ヶ月前にすでに大学を卒業していた	Non-Entailment
	過去	山田が去年家を建て終わったとき，鈴木はすでに大学を卒業していた	Non-Entailment
	現在	鈴木は大学を卒業している	Entailment
	未来 (副詞)	来年佐藤の就職先が決まるとき，鈴木は三ヶ月前に大学を卒業している	Non-Entailment
	未来	来年佐藤の就職先が決まるとき，鈴木はとっくに大学を卒業している	Entailment
現在		鈴木は大学を卒業している	
	過去 (副詞)	山田が去年家を建て終わったとき，鈴木は三ヶ月前にすでに大学を卒業していた	Non-Entailment
	過去	山田が去年家を建て終わったとき，鈴木はすでに大学を卒業していた	Non-Entailment
	現在 (副詞)	鈴木は三ヶ月前に大学を卒業した	Non-Entailment
	未来 (副詞)	来年佐藤の就職先が決まるとき，鈴木は三ヶ月前に大学を卒業している	Non-Entailment
	未来	来年佐藤の就職先が決まるとき，鈴木はとっくに大学を卒業している	Entailment
未来 (副詞)		来年佐藤の就職先が決まるとき，鈴木は三ヶ月前に大学を卒業している	
	過去 (副詞)	山田が去年家を建て終わったとき，鈴木は三ヶ月前にすでに大学を卒業していた	Non-Entailment
	過去	山田が去年家を建て終わったとき，鈴木はすでに大学を卒業していた	Non-Entailment
	現在 (副詞)	鈴木は三ヶ月前に大学を卒業した	Non-Entailment
	現在	鈴木は大学を卒業している	Non-Entailment
	未来	来年佐藤の就職先が決まるとき，鈴木はとっくに大学を卒業している	Entailment
未来		来年佐藤の就職先が決まるとき，鈴木はとっくに大学を卒業している	
	過去 (副詞)	山田が去年家を建て終わったとき，鈴木は三ヶ月前にすでに大学を卒業していた	Non-Entailment
	過去	山田が去年家を建て終わったとき，鈴木はすでに大学を卒業していた	Non-Entailment
	現在 (副詞)	鈴木は三ヶ月前に大学を卒業した	Non-Entailment
	現在	鈴木は大学を卒業している	Non-Entailment
	未来 (副詞)	来年佐藤の就職先が決まるとき，鈴木は三ヶ月前に大学を卒業している	Non-Entailment

表 7 時間関係と例文 (中国語) の一覧.

時間関係	文テンプレート	例文（中国語）
過去 (副詞)	[event-past] 的 时候，[np] 已 经[v][time] 了	田中上周搬家的 时候，山本已 经合格大学一周了
過去	[event-past] 的 时候，[np] 已 经[v] 了	田中上周搬家的 时候，山本已 经合格大学了
現在 (副詞)	[np] 已 经[v][time] 了	山本已 经合格大学一周了
現在	[np] 已 经[v] 了	山本已 经合格大学了
未来 (副詞)	[event-future] 的 时候，[np] 已 经[v][time] 了	佐藤下个月 换工作的 时候，山本已 经合格大学一周了
未来	[event-future] 的 时候，[np] 已 经[v] 了	佐藤下个月 换工作的 时候，山本已 经合格大学了

表 8 実験で使用したプロンプト (日本語&中国語)

指示: 前提と仮説の関係を entailment,non-entailment の中から回答してください. 説明は不要です.
制約：
- 前提から仮説が, 論理的知識や常識的知識を用いて導出可能である場合は entailment と出力
- 前提が成り立つとしても仮説が必ずしも成り立たない場合は non-entailment と出力
- 前提と仮説には, 時間的な成分を省略していない
- 前提と仮説の発話時を現在とする
前提: premise
仮説: hypothesis
答え:

指示: 从 entailment,non-entailment 中回答前提和假 设的 关系. 不需要 给出解 释.
限制：
- 如果能够通 过逻辑知 识或常 识性知 识从前提推 导出假 设，则输出 entailment.
- 如果前提成立无法保 证假 设成立,则输出 non-entailment.
- 前提和假 设中没有省略任何 时间成分.
- 前提和假 设的 发话时点为现在.
前提: premise
假 设: hypothesis
答案: