

記述文選択タスクによる 大規模視覚言語モデルのアモーダル補完能力の評価

綿引 周¹ 土井 智暉¹ 篠崎 大河^{2,1} 西田 知史^{3,4,5}

新川 拓哉⁶ 宮原 克典⁵ 谷中 瞳¹

¹ 東京大学 ² 慶應義塾大学 ³ 情報通信研究機構 ⁴ 大阪大学 ⁵ 北海道大学 ⁶ 神戸大学

amanew@g.ecc.u-tokyo.ac.jp snzktig@keio.jp s-nishida@nict.go.jp

{doi-tomoki701, hyanaka}@g.ecc.u-tokyo.ac.jp

niiitaku11@gmail.com kmiyahara@chain.hokudai.ac.jp

概要

本研究ではアモーダル補完 (amodal completion: AC) と呼ばれる現象に着目して大規模視覚言語モデル (Large vision language model; LVLM) の視覚的認識能力を評価した。ベンチマークの構築に上位オントロジーの一つである Basic Formal Ontology(BFO) を利用することでより体系的な評価を試みた。実験の結果、特に規模の大きな LVLM は全般的に人間に近い AC 能力をもつものの、人間と比較して、その能力には補完される対象の種類によってばらつきがあることが確認された。また LLaVA-NeXT モデルの一部については、基盤言語モデルのパラメータ数が大きくなっているにも関わらず正答率の低下が見られた。この結果は知覚動詞を適切に理解するという課題は基盤言語モデルのパラメータサイズを単純に増やすだけでは克服できない可能性を示している。

1 はじめに

大規模言語モデル開発の目標の一つは人間のマルチモーダルなタスクを補助するアシスタントを開発することにある。そのためには人間が自分の知覚経験を報告するテキストを LVLM が適切に理解できる必要がある。

アモーダル補完 (amodal completion: AC) は、遮蔽物や照明条件により対応する感覚刺激がないはずの物体や図形の一部が、それでもなお知覚されていると報告される現象である [1, 2, 3, 4]。例えば図 1 を提示された被験者は二つの長方形が「見える」と報告することが知られている [1]。しかし、一方の長方形の隠れた部分に対応する刺激は網膜に届いていない。そこで、その隠れた部分は知覚の中で「補完

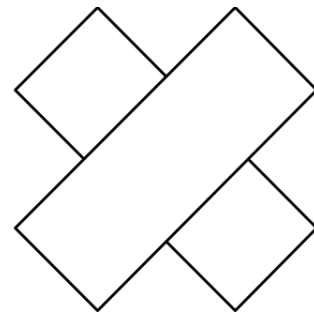


図 1 アモーダル補完の一例 (cf. [1], p. 143)

されているが、対応する感覚はないためその内容は「アモーダル」に補完されていると言われる [1]。AC は主体の側の何らかの能力を前提としていと考えられ、その能力を本稿では AC 能力と呼ぶ。

AC は日常生活で重要な役割を果たしており [5]、AC に関連する報告は人間が自分の経験を報告する種類のテキストに頻繁に現れると予想される。AC に関連する画像とテキストに適切に応答する能力は優れた LVLM の開発に不可欠であるだろう。

AC 能力は自動運転やロボットによる物体把握など、様々な応用上の意義を持つ。そのため近年、モデルがもつ AC 能力の評価が進んでいる [6]。しかし既存研究は主にモデルの画像処理能力の評価に焦点を当てており、AC に関連するテキストについての LVLM の推論能力は自明ではない。

そこで本研究では AC 現象を記述したテキストに関する LVLM の推論能力に着目する。具体的には多肢選択式タスクによるベンチマーク Visual Amodal Completion with Texts (VACT) を構築しモデルの推論能力を評価した。本ベンチマークでは [7] が提案した BFO を利用して補完される対象の種類を分類することでモデルの AC 能力の体系的な評価を試みた。

実験の結果 Claude 3.5 Sonnet や GPT-4o モデルは

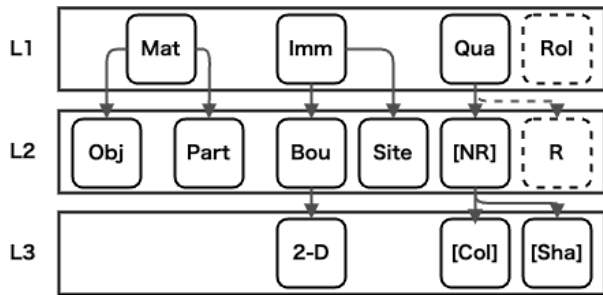


図2 BFO 持続者カテゴリのタクソノミーの一部。点線は本研究には関わらないカテゴリ。BFO にはこの図に含まれていない持続者のカテゴリもある。BFO 外のカテゴリにはブラケット ([]) をつけた。Mat: 物質的存在者; Imm: 非物質的存在者; Qua: 性質; Rol: 役割; Obj: オブジェクト; Part: オブジェクト部分; Bou: 境界; Site: 場所; NR: 非関係の性質; R: 関係の性質; 2-D: 2次元境界; Col: 色; Sha: 形。

AC に関連する多肢選択式タスクで概ね高い正答率を出したが、人間の安定した成績と比べると補完する対象のカテゴリによって得手/不得手があることがわかった。さらに LLaVA-NeXT モデルでは基盤となる大規模言語モデル (LLM) のパラメータサイズが大きいくほどベンチマークの成績が低下する現象が見られた。この結果は基盤となる LLM が大型であるほどマルチモーダルなタスクの遂行能力が向上するという一般的な傾向に反しており注目に値する。

2 関連研究

既存研究では様々な**画像処理**のタスクでモデルの AC 能力を評価するデータセットが構築されてきた [6]。しかし、いずれのタスクも、本研究で扱う AC を記述する**テキストについての LVLM の推論能力**を評価することはできない。

AC の種類によって、LVLM が得意とするものと苦手とするものがあるかもしれない。しかし、既存の研究で用いられてきた分類には問題がある。例えば [6] は近年の画像処理分野における AC タスクのうち重要なトピックとして形状補完、外観補完、順序知覚を挙げている。形態を補完するか、色を補完するかという観点で AC を分類する分類方法は他の研究にも見られる [4, 2]。[3] は補完されるものが「2D」か「3D」かの二種類で AC を分け、[8] はすべての補完は彼のいう「体積補完」であると論じている。しかし補完される対象のカテゴリは [6] が挙げる 3 種類以外にもある。また、先行研究では補完される対象のカテゴリの理解は自明なものとして前提され、より人工的でない画像に含まれる AC の事例を分類しようとする場合には困難が生じる。

3 提案手法

本研究では AC 現象についての記述文選択という多肢選択式タスクによるベンチマーク VACT を構築し、AC に関連するテキストについての LVLM の推論能力を評価した。既に現象学という哲学の分野で「地平意識」と呼ばれる経験の特徴に関する説明文を提示する多肢選択式タスクを人間の被験者に課した実験が存在する [9]。地平現象の少なくとも一部は AC と一致するため、[9] で用いられている設問と画像の一部は本研究でも利用可能である。

VACT では、Basic Formal Ontology (BFO) [7] に基づく AC の分類を提案する。BFO は異なる分野の科学研究を通して得られたデータの整理を一貫して、相互に参照可能な仕方で行うために開発された上位オントロジーである。オントロジーは実在世界に繰り返し現れる特徴を分類するカテゴリやカテゴリ間の関係を表すタクソノミーである。ここでタクソノミーとはカテゴリを表象する節 (node) とカテゴリ間のサブタイプないしサブクラス関係 (is_a 関係) を表す辺 (edge) を持つグラフである (図 2 参照)。

BFO の下では、最上位のカテゴリとして存在者 (存在する任意のもの) があり、そのサブタイプには持続者 (continuant) と生起者 (occurrent) がある。後者が時間的に展開する出来事や過程であるのに対して、前者は時間を通じてその同一性を維持する。時間的な出来事に関する AC も報告されている (「トンネル効果」と呼ばれる) [1] が、本研究で扱った AC の事例はすべて持続者にかかわるものだった。以下では主に持続者のみ、また持続者の中でもデータセット内の設問に関連するカテゴリのみ取り上げる。

持続者のサブタイプとして、性質の担い手となる独立的持続者とそうした独立的持続者の性質となり、それに「内在する」と言われる個別的依存的持続者がある。本研究ではこの階層のカテゴリを「第一層 (L1)」のカテゴリとして、第一層のカテゴリから「is_a」関係を下った子孫カテゴリを順に (is_a 関係の深さごとに) 「第二層 (L2)」 「第三層 (L3)」 のカテゴリと呼んでいる。各カテゴリの親子関係は図 2 の通りである。各カテゴリの説明は付録 A.2 を参照。

BFO の利点として 2 点が挙げられる。(1) 存在する任意のものが BFO のカテゴリのいずれかに分類されることが想定されている。それゆえ [6] の分類に当てはまらないカテゴリの対象にも対応できる。(2) 各カテゴリに明確な説明が与えられている。

画像	
選択肢	<p>正答 写真には写っていないが、グラスは、その前面と似た形状の背面をもつように見える</p> <p>誤答（誤補完） 写真には写っていないが、ストローの先はスプーンのような形状になっているように見える</p> <p>誤答（無補完） 写真には写っていないが、ストローの先はグラスの中まで続いているように見えない</p>

表 1 設問例

本研究では [6] と同様、補完対象の種類により AC タスクを分類する。分類にあたっては BFO を参照することで任意の種類を対象をより明確な基準に従って分類する。それにより LVLM の AC 能力についてきめ細かな評価が可能になると期待できる。BFO カテゴリとその定義に則って AC を部類する際に従った具体的な手続きは付録 A.3 を参照。

4 データセット構築

データ収集 本研究では [9]（全設問数 69 件）の一部を引き継ぎ、本研究の目的に沿って修正を加え、画像や設問を追加してデータセットを作成した。画像データはオンラインのロイヤリティフリーの画像サイトである Pixabay¹⁾ で収集し、追加の選択肢は著者たちが作成した。画像データはイラストと写真を凡そ同数用意した。問題文はすべて日本語である。

回答選択肢 二つある選択肢のうち一方は、AC 能力があれば容易に正解とわかるものだと想定される。もう一方の誤答となる選択肢としては「誤補完」（誤った内容を補完してしまっている記述）と「無補完」（補完が働いていない記述）という 2 パターンを用意した。図 1 に正答と誤答の 2 パターンの具体例を示した。

人手評価 人間が実際に正解文を選択できるのかを検証するために人手評価も行った。人手評価ではモデル評価のプロンプトと同じ指示を使用した。

一つの画像について誤答となる選択肢には二つのパターンを用意したが、正解の選択肢は一つしか作成しなかった。そのため問題群を二つのセットに分け、被験者一人につき各画像に関する設問は 1 パ

ターンしか提示されないようにした。

オンラインで 2025 年 12 月 6 日から同年 12 月 26 日の間に収集した 101 件（男性 63 名、女性 38 名；20 代 86 名、30 代 6 名、40 代 1 名）の回答結果から正答率が 50% を下回る設問は不適切な設問として除外した。その結果、合計 123 件の設問が有効な設問として残された（誤補完の選択肢を含む問題が 108 件、無補完の選択肢を含む問題が 107 件）。

5 検証

モデル LVLM の代表的な商用モデルとして GPT-4o²⁾ と Claude 3.5 Sonnet³⁾ を、Hugging Face Hub のオープンソースモデルとして複数のスケールの LLaVA-NeXT モデル [10, 11]（LLaVA-NeXT-34B, LLaVA-NeXT-72B, LLaVA-NeXT-110B）を評価対象とした。（以下、順に GPT-4o, Claude, LLaVA-34B, LLaVA-72B, LLaVA-110B と略記。）

LLaVA-NeXT シリーズはいずれも Vision Encoder + Connector + LLM という構成を採用している。LLaVA-72B と 110B の違いは学習に使用された GPU の種類と LLM のパラメータサイズのみである。

評価方法・指標 zero-shot で評価を行った。評価指標としては正答率を利用した。さらにテキストのみから尤もらしい回答を選択している可能性も検討するため、本来の画像の代わりにダミーとなる白紙の画像を提示した場合の正答率も求めた。ダミーでの実験も本来の画像での実験も同一のプロンプトを用いた。プロンプトの例は付録 A.1 を参照。

結果 補完される対象の種類毎の LVLM と人間の

1) <https://pixabay.com/>

2) <https://openai.com/index/GPT-4o-system-card/>

3) <https://assets.anthropic.com/m/61e7d27f8c8f5919/original/Claude-3-Model-Card.pdf>

表 2 実験結果 1. アモーダル補完される対象のカテゴリごとの各モデルの正答率. 括弧内の数値は白紙画像を提示したダミー問題の正答率. ダミー問題の正答率が本来の問題の正答率より大きい箇所を灰色で示した. 各カテゴリ一名の説明は図 2 を参照. 人手評価の値は平均の正答率.

L1	Mat		Imm		Qua	
	Obj	Part	Bou	Site	[NR]	
L2					[Col]	[Sha]
L3			2-D			
設問数	16	39	54	18	36	12
GPT	88 (75)	77 (74)	94 (72)	100 (100)	86 (64)	92 (83)
Claude	100 (81)	87 (90)	93 (93)	94 (100)	83 (89)	100 (83)
LLaVA -34B	69 (63)	46 (43)	56 (50)	72 (61)	56 (50)	75 (67)
LLaVA -72B	63 (50)	70 (43)	57 (48)	100 (100)	56 (31)	67 (75)
LLaVA -110B	46 (63)	57 (48)	57 (52)	89 (83)	53 (50)	67 (67)
人間	95	94	94	92	92	95

正答率を表 2 に示した. 商用モデルは人間に近い成績を残しているが, BFO を取り入れた分類により各モデルが苦手とするカテゴリがあることが鮮明になった. GPT-4o はオブジェクト部分の補完を苦手とし, Claude は色の補完を苦手とする. LLaVA モデルはいずれも境界と色補完を苦手としている.

モデルはダミー問題でもかなりの正答率に達している. 中にはダミー画像の方が正答率が高いカテゴリも存在し, Claude では半数のカテゴリでダミー画像の方が良い成績を残している.

LLaVA-110B の正答率がそれより小規模な LLaVA-72B よりも低下している点も注目に値する. 後者が正答しながら前者が誤答している問題はどのカテゴリにも存在するが, 各カテゴリの全設問数を考慮してもオブジェクト部分 (10 問) の誤答率が高い.

誤答の選択肢の種類による誤答数の変化も調べた (表 3). 誤答した全問題数に対する無補完での誤答の割合が GPT-4o 以外のすべてのモデルで, ダミー画像よりも本来の画像を提示した場合に高くなった.

考察 今回の実験結果について, 以下の説明が考えられる. (1) 今回用意した AC を報告する正解文は誤りの選択肢に比べてより自然な言語表現になっているため, 選択肢だけからは正解の文を選びやすい. (2) 基盤とする LLM のパラメータサイズがより大きい LVLM ほどマルチモーダルベンチマークにおいて常に優れた性能を発揮するが [12], マルチモーダルなタスクの遂行能力が高いほど LVLM の応答は画像

表 3 実験結果 2. 誤りの選択肢のパターンによる誤答数. 誤答した全問題の中で誤補完の割合はダミー問題よりも本来の画像を提示した場合に増加することが見て取れる. GPT-4o に誤補完数の割合はいずれも 75% だった.

	誤補完	無補完	全誤答数中の 無補完の割合
Claude (dummy)	7 (12)	9 (5)	56% (29%)
LLaVA-72B (dummy)	29 (44)	35 (43)	57% (49%)
LLaVA-110B (dummy)	36 (42)	35 (5)	49% (45%)

の内容を反映しやすくなると考えられる. しかし, (3) 対象が直接視野に現れていないが, それでもあるように「見える」という表現が成り立つ知覚動詞の用法をモデルが十分に把握していない. それゆえ本来の画像が提示された場合, 言及されている対象はその画像に含まれないために, 選択肢だけからは正解として選びやすい記述が誤答と判断されやすくなる. この点は, ほぼ全てのモデルで, ダミー問題よりも本来の問題で無補完の選択肢を選んでしまう割合が増えていることから支持されうる (表 3).

しかし, AC を含む経験の報告を理解するためには上述の知覚動詞の用法を LVLM が適切に扱える必要がある [13]. Claude や LLaVA は「見える」のこの用法を識別できておらず, だからこそ LLaVA は基盤 LLM のサイズが大きくなったときに正答率が下がったと考えられる. 一方, GPT-4o はすでにこの課題を克服しているために本来の問題で, ダミー問題よりも高い正答率を残している可能性がある.

6 おわりに

本研究では, LVLM のアモーダル補完 (AC) 能力を AC に関連する推論能力に着目して評価した. その際, BFO を参照することで体系的な評価を試みた. 実験の結果, LVLM は全般的に人間に近い能力をもつものの, 人間と比較して, その能力には補完される対象の種類により得手/不得手があることが確認された. またダミー画像を用いた実験により, LLaVA や Claude では AC に関連するテキストを理解するために必要な知覚動詞の適切な理解が欠けている可能性が示された. 今後の展望として, モデルが知覚動詞の一部の用法を理解できていないという仮説について, さらなる検証を進めたい.

謝辞

本研究は JSPS 科研費学術変革領域研究 (B) 「ナラティブ意識学」JP24H00809 の支援を受けたものである。

参考文献

- [1] A Michotte. Amodal completion of perceptual structures. In Georges Thinés, Alan Costall, and George Butterworth, editors, **Michotte's experimental phenomenology of perception**, pp. 140–167. Lawrence Erlbaum, Associates, 1991.
- [2] Luiz Pessoa, Evan Thompson, and Alva Noë. Filling-in: One or many? **Behav. Brain Sci.**, Vol. 24, No. 6, pp. 1137–1139, December 2001.
- [3] Rob van Lier and Walter Gerbino. **Perceptual completions**. Oxford University Press, April 2015.
- [4] Walter Gerbino. Amodal completion revisited. **Iperception**, Vol. 11, No. 4, p. 2041669520937323, July 2020.
- [5] Bence Nanay. The importance of amodal completion in everyday perception. **i-Perception**, Vol. 9, No. 4, pp. 1–16, 2018.
- [6] Jiayang Ao, Qihong Ke, and Krista A Ehinger. Image amodal completion: A survey. July 2022.
- [7] Robert Arp, Barry Smith, and Andrew D Spear. **Building Ontologies with Basic Formal Ontology**. The MIT Press. MIT Press, London, England, July 2015.
- [8] P U Tse. Volume completion. **Cogn. Psychol.**, Vol. 39, No. 1, pp. 37–68, August 1999.
- [9] Satoshi Nishida, Hiro Taiyo Hamada, Takuya Niikawa, and Katsunori Miyahara. Neural correlates of phenomenological attitude toward perceptual experience. **bioRxiv**, 2024.
- [10] Haotian Liu, Chunyuan Li, Yuheng Li, and Yong Jae Lee. Improved baselines with visual instruction tuning, 2023.
- [11] Bo Li, Kaichen Zhang, Hao Zhang, Dong Guo, Renrui Zhang, Feng Li, Yuanhan Zhang, Ziwei Liu, and Chunyuan Li. Llava-next: Stronger llms supercharge multimodal capabilities in the wild, May 2024.
- [12] Bo Li, Kaichen Zhang, Hao Zhang, Dong Guo, Renrui Zhang, Feng Li, Yuanhan Zhang, Ziwei Liu, and Chunyuan Li. LLaVA-NeXT: Stronger LLMs supercharge multimodal capabilities in the wild, May 2024.
- [13] 彩子志波. 知覚動詞「見える」の推定構文への拡がり: 共時態における文法化. 名古屋大学人文学研究論集, Vol. 6, pp. 57–78, 2023.
- [14] Bjoern Peters, Chris Mungall, David Osumi-Sutherland, Jie Zheng, Mathias Brochhausen, Barry Smith, Alan Ruttenberg, Albert Goldfain, Bill Duncan, Fabian Neuhaus, James A. Overton, Janna Hastings, Jonathan Bona, Larry Hunter, Leonard Jacuzzo, Ludger Jansen, Mark Ressler, Mauricio Almeida, Melanie Courtot, Pierre Grenon, Randall Dipert, Ron Rudnicki, Stefan Schulz, Thomas Bittner, Werner Ceusters, Yongqun “Oliver” He, Holger Stenzhorn, Selja Seppälä. Ontobee: BFO. <https://ontobee.org/ontology/BFO>. Accessed: 2025-1-3.

A 付録

A.1 実験に使用したプロンプト

— 実験に使用したプロンプト —

指示: 画像と、その画像を見ているときの体験内容を説明する文が2つ出てきます。一方は正しく、もう一方は誤っています。説明文1と説明文2のどちらが適切か、説明文の番号で、また番号のみで回答してください。

説明文1: 写真には写っていないが、グラスは、その前面と似た形状の背面をもつように見える

説明文2: 写真には写っていないが、ストローの先はスプーンのような形状になっているように見える

選択肢を提示する順序と選択肢の記号によるバイアスの影響を回避するため説明文1が正解となる設問と説明文2が正解となる設問の数が半々に近い数になるようランダムに説明文の順序を入れ替えた。

A.2 BFO カテゴリの説明

以下に本研究に登場したBFOカテゴリの定義と適宜、そのカテゴリの具体例とそのカテゴリではない存在者の例(「非例」)を挙げる⁴⁾。紙幅の都合上定義は多少簡略化している。正確な定義は[7, 14]を参照せよ。

A.2.1 第1層

物質的存在者 Material entity; Mat [定義] 物質を部分としてもつ独立的持続者。物質的存在者は三次元に空間的に広がっていて、どれほど短くとも一定の時間間隔を通じて存在し続ける。[例] 人間などの有機体、人間の体についている腕。[非例] 有機体の表面、運河(の窪み)

非物質的存在者 Immaterial entity; Imm [定義] 物質的存在者を部分として含まない独立的持続者。

性質 Quality; Qua [定義] もしある存在者に内在する場合には、その存在者において完全に顕現し、実現される個的依存的持続者。[例] 腎臓の質量、血液のこの部分の色、手の形。[非例] この砂糖の水溶性、心臓の機能。

A.2.2 第2層

対象 Object; Obj [定義] 以下3つを満たす物質的存在者のことである。(1) 三次元に空間的に延長している；(2) 因果的に統一されている。すなわち、もし当のオブジェクトの一部が空間内を移動するならば他の部分も移動する確率が高いという仕方、その諸部分が結合している；(3) 最大自己連結的(maximally self-connected)である⁵⁾。[例] 有機体や単細胞、宇宙船、惑星

オブジェクト部分 Fiat object part; Part [定義] 何らかのより大きなオブジェクトの真部分だが物理的な断続により、このオブジェクトの他の部分から区別されているわけではない物質的存在者。[例] あなたの胴。[非例]

4) 正確にはBFOカテゴリに与えられるのは定義というよりも説明(elucidation)である[7, 14]。

5) オブジェクトが最大限自己連結的であるとは直観的にはオブジェクトの一部から連続的な線を追ってそのオブジェクトの外に出ることなくそのオブジェクトの他の任意の部分へ移動できることを意味する。

血管内の赤血球や鰭の缶詰の中の個々の鰭⁶⁾

境界 Fiat object boundary; Bau [定義] 0, 1, 2次元の非物質的存在者で空間領域を部分として含まない。直観的には当の対象がその周囲と出会うまさにその場所に存在する、物質的存在者の境界。本研究では2次元の境界のみ取り上げた。[例] 地球の表面や細胞膜の表面

場所 site [定義] (1) 物質的存在者により(部分的にか全体的に)境界づけられているか、(2) (1)を満たす3次元の非物質的部分であるかのいずれかである三次元の非物質的存在者である。[例] 鼻腔、血管(血液が流れる空洞)、スエズ運河(溝の部分)。

非関係的性質 Non-relational quality; NR [定義] BFOには関係的性質というカテゴリが性質のサブタイプとして用意されている。関係的性質は二つ以上の独立的持続者に内在する性質である。非関係的性質は正確にはBFOカテゴリではないが、一つの独立的持続者に内在する性質として自然に定義できる。[例] 色、形。[非例] 婚姻関係(関係的性質の例)

A.2.3 第3層

2次元境界 Two-dimensional continuant fiat boundary; 2-D 自己連結的(self-connected)でその位置が何らかの物質的存在者と相対的に定義される表面。[例] 地球の表面 [非例] 土地の境界(一次元境界の例)

色 Color; Col, 形 Shape; Sha この二つのカテゴリはBFOカテゴリである性質のサブタイプだが、BFOの一部ではない。それでもAC補完を分類する既存の研究でしばしば言及されることから今回はこのふたつのカテゴリを結果に含めた。同様の仕方では他のカテゴリ(オブジェクトや場所など)の配下に特定分野の分類体系を接続することができる。

A.3 AC分類ガイドライン

1. 正解の記述で言及されている対象のうち画像内でアモータルに補完されている対象 x を特定する。
2. 正解の記述と各BFOカテゴリ ϕ の定義から、またそれのみから「 x は ϕ である」が含意されるかどうかを判断する。その際、自分が当のACを経験する場面を想像して判断しないように注意する。
3. 「 x は ϕ である」が含意される場合には、 ϕ カテゴリの対象が補完されていると考える。

図1の設問を例に説明する。正解の記述で補完される対象として言及されているのはグラスの背面である。グラスの背面はグラスがその周囲と出会うまさにその場所に存在する表面である。それゆえ2次元境界の説明と正解の記述から、補完されるその対象 x (すなわち写真のグラスの背面)について「 x は2次元境界である」ことが含意されると判断できる。もしかするとこの画像を見ているあなたは何らかの色を持った表面がグラスの背面に続いているようにも見えるかもしれない。しかしグラスの背面がある特定の色であることは**正解の記述からは含意されない**。それゆえ選択肢を正解として選ぶために要求されるAC能力は色ではなく2次元境界の補完能力であると判断する。

6) 「fiat」という語は、問題の境界が、例えば地図の描き手や理論家の側での決定を反映しているという事実に向けられるものである。