Hansatic

技術ホワイトペーパー

ニューロシンボリック積載計画

実顧客の積載で検証した積付けエンジンと8つの最先端AIモデル

2026年11月

hansatic.com/ja/neuro-symbolic-load-planning

すべてのプランを3Dで確認できる関連ページ: hansatic.com/ja/neuro-symbolic-load-planning

概要

確定的積載エンジンと生成AIを組み合わせている。組み合わせる前に、それぞれを単独で計測した。

- **データ:** 実顧客の積載計画ケース20件、採点済みラン26回。海上コンテナ、カーテンサイダー、53 ft (16.2 m) ドライバン、AKE航空コンテナ、連結パレット構成を含む。
- **参加者:** Hansaticの本番積載エンジンと8つの主要フロンティアLLM。各モデルはツールなしで、1つのJSON回答ですべてのアイテムの位置と向きを示す。
- **評価者:** 1つの確定的バリデータが全参加者を同じ方法で採点する。チェック項目：
  - 形状；
  - 95%の底面支持；
  - 積み重ね・壊れ物制限；
  - 向きフラグ；
  - 細長い積み重ね；
  - 積載重量と軸重；
  - 前後方向の重心範囲。

結果：

- **エンジン**は26回中26回で合法。各ランの最良合法モデル回答と比べ、18回で優位（13回は明確に、5回は僅差）、8回で同点。劣ることはない。
- **最強モデル、** GPT-5.5は26回中25回で合法。最大253個のマニフェストを含む。
- **8モデル中4モデル**は3分の1未満のランでのみ合法。
- **モデルの得意分野。** 最良のモデル回答はエンジンになかった2つの配置戦略を使った。どちらも今は確定的エンジン戦略だ。

本日現在、エンジンの方が優れた計画を立てられると結論づける。モデルの創造性は取り込む価値がある。そのためのアーキテクチャがニューロシンボリックだ：

- モデルが戦略を提案し、
- エンジンがすべての計画を構築し、
- バリデータが合法な改善のみを保持する。

ケース、バリデータ、プロンプト、スコアリングコードを公開する。

1. ベンチマーク設計

1.1 ケース

各ケースはHansaticアプリの実顧客積載を1:1でコピーしたものです。寸法、重量、数量、フラグ、車両はそのまま。品目IDは置き換え、貨物名は削除しています。

セット	ケース	ラン	車両	個数
標準 (H01-H12)	12	16	20' 1台、40' ドライ 2台、40' HC 4台、45' HC 1台、13.6 m (44.6 ft) ウイング車 4台	14-83
大型 (BV1-BV6)	6	8	40' HC 3台、40' ドライ 1台、13.6 m (44.6 ft) ウイング車 1台、16.2 m (53 ft) ドライバン 1台	103-253
特殊 (U01, P01)	2	2	傾斜壁付きAKE (LD3)、1200 x 1000 mm (47 × 39 in) パレット	50と40

ランと目的:

- **トラック**は積載メートル（LDM）とパレットポジションの2軸で評価する。これで標準12ケースから16ラン、大型6ケースから8ランとなる。
- **BV5とBV6**は全量が入らない。そこではより多くの品目を置くことが主目的だ。
- **U01とP01**には既知の最適値がある:
  - U01: 50個中24個。傾斜壁のためこれ以上入らない;
  - P01: 40袋中40袋、インターロック率100%。

ケースの制約:

- 段積み不可および天積みのみ（フラジール）の品目;
- 品目ごとの段積み重量制限;
- 天地指定および傾斜禁止のフラグ;
- 長尺バー（BV1、コンテナ長まで）;
- 重量が突出した単一品目（H02）;
- パレットとカートンの混載（BV3）。

1.2 バリデータ

バリデータ（v2.3.0）はすべての計画を同じ方法で検査する。すべてのハードチェックに合格した場合にのみ、計画は合法となる。

チェック	ルール
範囲	すべての品目が貨物スペース内にあること（AKEの傾斜面を含む）

チェック	ルール
重複	2つの品目が体積を共有しないこと
支持	各品目の底面の95%以上が床または下の品目の上にあること
段積み	段積み不可の品目は何も載せない。天積みのみの品目は何も載せず最上部に置く。品目ごとの段積み重量制限を守ること
向き	天地指定および傾斜禁止フラグを守ること
細長い山	自立スタックにおける高さと底面の比率に制限あり
積載量	総質量が車両制限内であること
車軸（トラック）	前後の軸重が制限内であること（品目の重心から計算）
CoG（コンテナ）	長手方向の重心が許容バンド内にあること
複数停車	停車がある場合は停車順に後入れ先出し

転倒リスクは計算・報告されるが、採点対象外。

1.3 ランキング

ランごとに適用:

1. 合法な計画は違法な計画に勝る。
2. 置いた品目が多い方が勝ち。
3. ランの目的で決める:
    - より短い積載（LDM）;
    - またはパレットポジションが少ないこと;
    - またはP01ではインターロック率が高いこと、次に高さが低いこと。
4. パレットポジションが同数の場合はLDMで決着をつける。

パレットポジションを数えるカウンターは全参加者で1つ。Hansaticの本番カウンターだ。

マージン。 勝ちを次のように呼ぶ:

- LDMで2%以上、パレットポジションで1つ以上、または置いた品目が1つ以上多い場合を**明確**とする;
- それ以外は**僅差**とする。

エンジンは時間予算内で動くため、単一ランでの2%未満の差はラン間のばらつきの可能性がある（6.2節）。

1.4 参加者

モデル。 8ラボそれぞれの主流フロンティアモデル:

- GPT-5.5 (OpenAI)
- Claude Opus 5.5 (Anthropic)
- Grok 4.7 (xAI)
- Gemini 4 Argon (Google)
- Kimi K3.1 Pro (Moonshot AI)
- DeepSeek V4 Pro (DeepSeek)
- Qwen3p8 (Alibaba)
- Magistral Medium (Mistral AI)

設定:

- **ラン:** ケースごとに1回。
- **出力制限:** 64kトークン、またはプロバイダー側の低い最大値（Gemini: 32k）。
- **バッチ処理:** 提供される場合はプロバイダーのバッチAPIを使用。
- **回答収集日:** 2026-11-11。
- **総コスト:** 約$64。

エンジン。 Hansaticの本番パッカー、2026-11-20版（コミット 9f66537c）。設定:

- 本番設定: 4ワーカーと本番時間予算、1 CPUコアに固定;
- ケースごとに1ラン、ベストオブなし;
- 独自の60%支持ルールで梱包し、他のモデルと同様に95%で評価する。

1.5 プロンプト

- 車両種別ごとに1プロンプト: コンテナ、トラック、ULD、パレット。
- 各プロンプトには、その種別にバリデータが適用する正確なルール、目的、出力スキーマ（品目ID、位置、向き）を記載する。
- 解答済み配置の例は示さない。

完全なテンプレートは付録Aにある。

2. 結果

2.1 モデル別、全26ラン

モデル	合法	エンジンと同等	エンジンに劣る
GPT-5.5	25	4	22
Opus 5.5	20	6	20
Grok 4.7	20	4	22
Kimi K3.1 Pro	12	5	21
DeepSeek V4 Pro	7	3	23
Gemini 4 Argon	5	2	24
Qwen3p8	4	3	23
Magistral Medium	2	0	26

どのモデルもエンジンに勝るランはなし。

標準セット（16ラン）

モデル	合法	同等	劣る
GPT-5.5	15	3	13
Opus 5.5	13	4	12
Grok 4.7	10	2	14
Kimi K3.1 Pro	8	3	13
DeepSeek V4 Pro	4	2	14
Gemini 4 Argon	4	1	15
Qwen3p8	2	1	15
Magistral Medium	1	0	16

大規模セット（8ラン） どのモデルも大規模ランでエンジンと同等の結果はなし。

モデル	合法
GPT-5.5	8
Grok 4.7	8
Opus 5.5	5

モデル	合法
Kimi K3.1 Pro	2
DeepSeek V4 Pro	1
Magistral Medium	1
Gemini 4 Argon	0
Qwen3p8	0

特殊セット（2ラン）

モデル	合法	最適解に到達
Opus 5.5	2	2
Grok 4.7	2	2
Kimi K3.1 Pro	2	2
Qwen3p8	2	2
GPT-5.5	2	1
DeepSeek V4 Pro	2	1
Gemini 4 Argon	1	1
Magistral Medium	0	0

2.2 エンジンと最良の合法モデル回答の比較

セット	明らかに優位	僅差で優位	同等	劣る
標準（16）	7	3	6	0
大規模（8）	6	2	0	0
特殊（2）	0	0	2	0
全体（26）	**13**	**5**	**8**	**0**

標準セット 各セルは積載アイテム数、続いて積載長またはパレット位置を示す。

ラン	エンジン	最良の合法モデル	マージン
H01	26/27, 11.69 m (38.4 ft)	Opus 26, 11.74 m (38.5 ft)	僅差 (0.4%)
H02	24/29, 5.78 m (19.0 ft)	GPT 18, 5.76 m (18.9 ft)	明らか (+6個積載)
H03	31/32, 10.51 m (34.5 ft)	合法なし	明らか
H04	39, 10.44 m (34.3 ft)	Opus 39, 10.44 m (34.3 ft)	同等

ラン	エンジン	最良の合法モデル	マージン
H05	49, 12.40 m (40.7 ft)	Opus 49, 12.40 m (40.7 ft)	同等
H06	83, 10.65 m (35.0 ft)	Grok 83, 11.41 m (37.4 ft)	明らか (6.6%)
H07	45, 9.60 m (31.5 ft)	GPT 45, 9.60 m (31.5 ft)	同等
H08	52, 5.92 m (19.4 ft)	Kimi 52, 6.12 m (20.1 ft)	明らか (3.3%)
H09 LDM	29/30, 13.36 m (43.8 ft)	GPT 29, 13.56 m (44.5 ft)	僅差 (1.5%)
H09 パレット	29/30, 29 ポジション	GPT 29, 29 ポジション	僅差 (LDMタイブレーク)
H10 LDM	26, 12.20 m (40.0 ft)	DeepSeek, 同等	同等
H10 パレット	26, 31 ポジション	DeepSeek, 同等	同等
H11 LDM	23, 12.00 m (39.4 ft)	Opus 23, 12.20 m (40.0 ft)	僅差 (1.6%)
H11 パレット	23, 30 ポジション	Kimi 23, 32 ポジション	明らか (2ポジション)
H12 LDM	13/14, 11.00 m (36.1 ft)	Gemini 13, 11.00 m (36.1 ft)	同等
H12 パレット	13/14, 25 ポジション	Kimi 13, 28 ポジション	明らか (3ポジション)

大規模セット

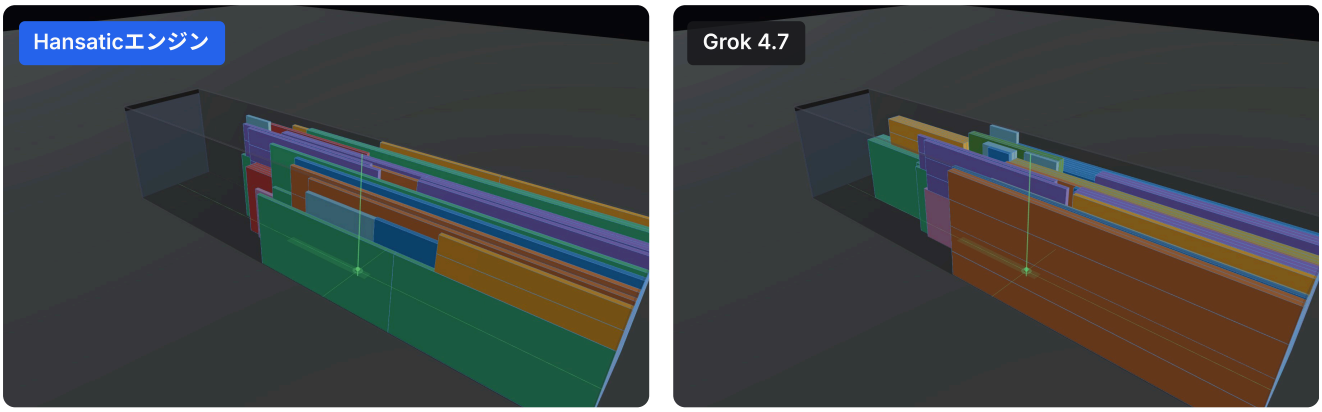
ラン	エンジン	最良の合法モデル	マージン
BV1 (40' HC, 103本)	103, 9.17 m (30.1 ft)	Grok 103, 10.66 m (35.0 ft)	明らか (14.0%)
BV2 LDM (カーテンサイダー, 153)	153, 11.52 m (37.8 ft)	Opus 153, 11.78 m (38.6 ft)	明らか (2.2%)
BV2 パレット	153, 29 ポジション	Opus 153, 29 ポジション	僅差 (LDMタイブレーク)
BV3 LDM (53 ft バン, 119)	119, 6.83 m (22.4 ft)	Opus 119, 6.99 m (22.9 ft)	明らか (2.2%)
BV3 パレット	119, 17 ポジション	Opus 119, 18 ポジション	明らか (1ポジション)
BV4 (145, 上部のみ割れ物)	145, 8.57 m (28.1 ft)	Grok 145, 8.85 m (29.0 ft)	明らか (3.1%)
BV5 (136, 縦置きのみ)	127, 12.00 m (39.4 ft)	Opus 124, 11.77 m (38.6 ft)	明らか (+3個積載)
BV6 (253, 容積の95%)	244, 11.98 m (39.3 ft)	GPT 238, 11.97 m (39.3 ft)	明らか (+6個積載)

特殊セット

ラン	エンジン	モデル
U01 (AKE)	24/50, 最適解	Opus、Grok、Kimi、Qwen、Geminiも24に到達；GPTとDeepSeekは18に到達
P01 (パレット)	40/40, 100%インターロック, 最適解	すべての合法モデル回答が40/40の最適解に到達

2.3 マージンが実際の積載に与える影響

- **BV1:** 1.49 m (4.9 ft) 短縮、40フィートコンテナの長さの約12%。
- **H12 パレット:** 3ポジション少ない、13.6 m (44.6 ft) カーテンサイダーのデッキの約9%。
- **BV6:** 積載しきれない荷物に6個多く積載、別の車両に回す荷物が6個少ない。



BV1: Hansaticエンジン 9.17 m (30.1 ft)、左；最良の合法モデル回答（Grok 4.7）10.66 m (35.0 ft)、右。

2.4 コストと時間

	エンジン	モデル
計画あたりの時間	無視できる	比較せず：バッチAPIはリクエストをキューイング
計画あたりのコスト	無視できる	$0.10-0.80

3. 失敗分析

208回のモデル実行（8モデル x 26回）のうち、95件の計画が適法です。残りの113件は3つの異なる形で失敗しています。

モデル	適法	回答なし（出力制限）	プロバイダーエラー	ルール違反の計画
GPT-5.5	25	1	0	0
Opus 5.5	20	6	0	0
Grok 4.7	20	0	4	2
Kimi K3.1 Pro	12	14	0	0
DeepSeek V4 Pro	7	19	0	0
Gemini 4 Argon	5	2	0	19
Qwen3p8	4	22	0	0
Magistral Medium	2	1	0	23
全体	**95**	**65**	**4**	**44**

3.1 回答なし：モデルが考える余地を使い果たす

65回の実行で、モデルは出力制限に達するまで推論しましたが、計画を書きませんでした。さらに1件の回答（Magistral、H03）は正常に終了しましたが、有効なJSONではありませんでした。これはここに含めています。

- これは最も一般的な失敗で、8モデル中6モデルに影響しています。
- Kimi、DeepSeek、Qwenはこの理由でほとんどの実行を失います。彼らの適法な計画は主に小規模なケースにあります。
- Opusはこの理由で6回の実行を失います。その中にはBV1、BV4、BV6があり、最大級のマニフェスト3件です。

生産環境のプランナーにとって、回答がないことは違法な計画と同じ失敗です。ドックには依然として計画が必要だからです。

3.2 プロバイダーエラー

プロバイダーの支出上限に達したため、Grokは4回の実行でHTTP 403を返しました。これらの実行は再送されず、失敗として採点されています。

3.3 ルールに違反する計画

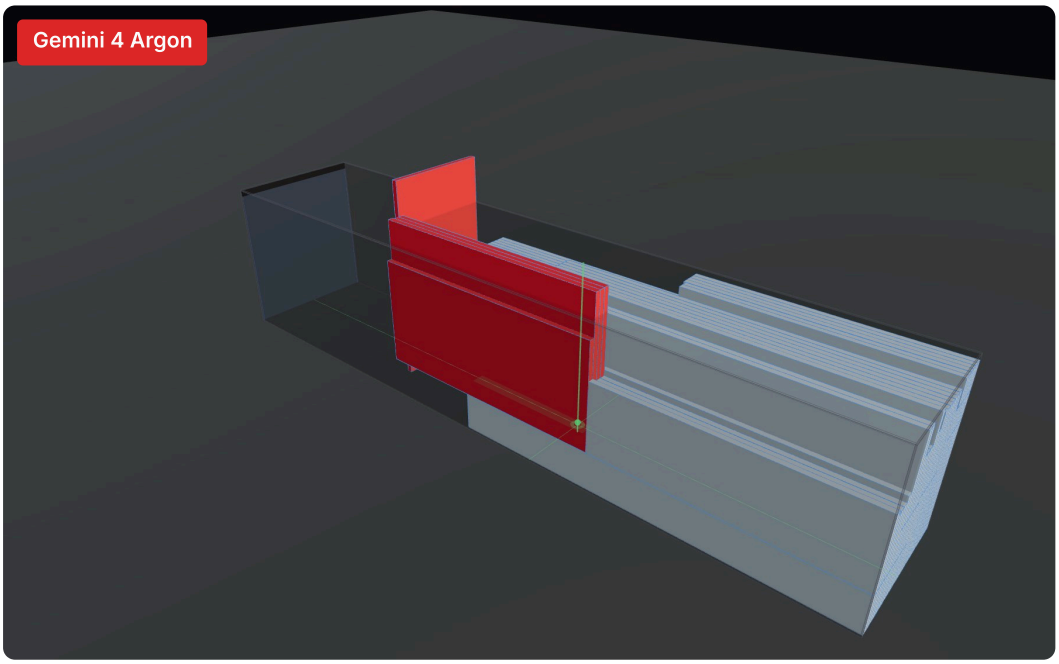
44件の回答は、少なくとも1つのハードルールに違反する完全な計画です。ほぼすべてが2つのモデルによるものです。Gemini（19件）とMagistral（23件）。Grokは2件です。

これら44件の計画のうち、各種類の違反を少なくとも1件含む計画の数：

違反したルール	計画数
品目が積載スペース外にある	27
品目が重なり合っている	27
浮いている品目（支持率95%未満）	21
許可された向きに合致しない品目寸法	17
段積み不可または割れ物の上への積載、または段積み制限の超過	10
屋根より上への積み重ね	9
重心が許容範囲外	7
軸重超過	4
傾け禁止品目が傾けられている	1
AKE輪郭の外側	1

Grokの2件の違法な計画はそれぞれ1つのルールに違反しています。ある実行では潰れた品目、別の実行では重心です。

これらの計画は3D表示ではもっともらしく見えます。欠陥はしばしば数個の品目で、センチメートル単位で浮いたり重なったりしています。完全な幾何学的チェックでのみそれらを発見できます。



BV1、Gemini 4 Argon：103品目中77品目を配置。ハイライト：浮いている、重なっている、または積載スペースからはみ出している品目。

3.4 適法だが、遅れている

- **不完全。** BV5とBV6では、最良の適法なモデル回答はエンジンより3個および6個少ない品目を配置します。
- **長い。** すべての品目が収まる実行において、最良の適法なモデル回答は24回のトラックおよびコンテナ実行のうち13回でエンジンのものより長くなっています。最大の差：
  - BV1で1.49 m (4.9 ft)；
  - H06で0.76 m (2.5 ft)。

3.5 モデルが優れている点

最も優れた回答は、経験豊富な作業員が使用する戦略を用いています：

- **重い木箱を前壁に当てる。**
- **段積み不可品目の上の高さを利用する。** BV3で、Opusはカートンを横にして高さ559 mm (22 in)の床として敷き、その上に19台の段積み不可の家電製品を立てます。

エンジンにはどちらの戦略もありませんでした。両方とも現在は決定論的エンジンの戦略です。これはニューロシンボリックループが捉えることを目的とした創造性です（セクション5）。

4. ツールアーム

第2のアームでは、モデルは1つのツール `pack_cargo` を通じてエンジンを呼び出し、その後、任意の計画で応答できる。

ほとんどのモデルはマニフェストをツール呼び出しにコピーし、エンジンの計画をそのまま返す。失敗はコピーエラーである：

- Magistralは寸法またはフラグが誤った100個のアイテムを送信した;
- KimiはH10で誤った呼び出しを行った;
- Qwenは4ランで失敗した。

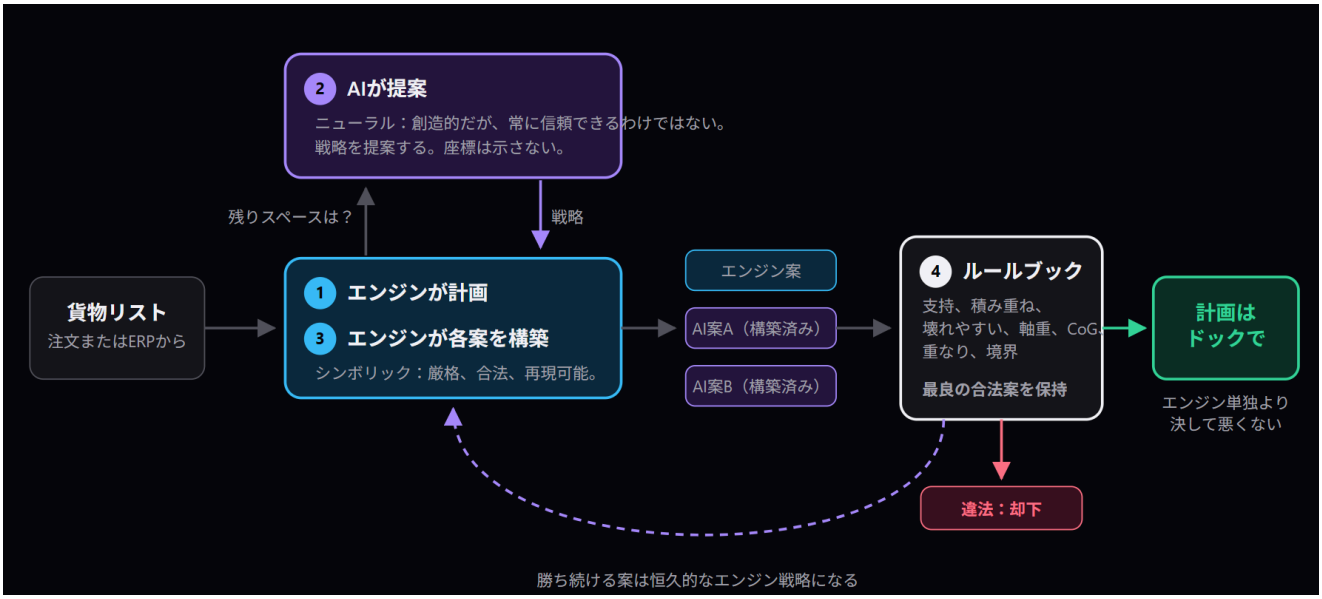
示唆: あらゆる統合システムにおいて、エンジンはモデルを介さず、ソースシステムから直接マニフェストを読み取らなければならない。

5. ニューロシンボリック・アーキテクチャ

5.1 二種類の知性

	パッキングエンジン（シンボリック）	生成AI（ニューラル）
ルール	エンコードされた全ルールに毎回従う	与えられたルールを破る
再現性	同じ積荷には同じ計画	同じ積荷でも異なる計画になりうる
監査	すべての配置に理由がある	監査が困難
戦略の幅	保有する戦略に限定	新しい戦略を見つける
規模	253個を処理できる	大規模マニフェストでは回答不能なモデルが多い

5.2 ループ



1. **ベースライン。** エンジンがマニフェストから計画を立てる。マニフェストはソースシステムから直接読み込む。
2. **トリガー。** モデルを呼ぶのはベースラインに余地があるときだけ。未配置の貨物、低充填、クリアランス目標の未達。
3. **提案。** モデルは戦略パラメータを返す。座標は返さない：
   - ゾーン；
   - 順序；
   - 向きの優先順位；
   - 床とキャップの構造。
4. **実行。** エンジンは提案ごとに1つの計画を構築する。

5. **判定。** バリデータが最良の適法計画を選ぶ。ベースラインは常に候補に入る。
6. **蒸留。** 繰り返し勝った提案はエンジンの恒久戦略となる。前壁戦略やカートン床戦略のように。

5.3 構造上の特性

- **エンジン単体より悪くならない。** ベースラインが常に候補だから。
- **違法な計画がドックに届かない。** バリデータが全候補を判定するから。
- **転記リスクなし。** モデルはマニフェストも座標も触らないから（セクション4）。
- **出力制限による失敗なし。** 提案は数個のパラメータで、253個の計画ではないから（セクション3.1）。
- **AIコストに上限あり。** モデルはトリガー時のみ動くから。

このループを現在構築中。次回の版では計画ごとのコストと時間を含め、独立したエントリーとして採点する。

6. 範囲と限界

6.1 バリデーターの範囲

合法とは、v2.3.0に組み込まれたルールに基づく合法性を指す。バリデーターは、マルチストップ時の後入れ先出し以外の積載順序をチェックしない。

6.2 ばらつき

各参加者はケースごとに1回実行する。

- **エンジン**は時間制限内で戦略を競わせる。マシン負荷により、先に完了する戦略が変わることがある。H02では実行間で5.63-5.79 m (18.5-19.0 ft)を計測した。
- **モデル**は1回だけサンプリングする。

単一実行で2%未満の差は「僅差」と報告する。

6.3 モデルティアと出力制限

- 各ラボの主要フロンティアモデルを実行した。プレミアムティア（Anthropic Fable、OpenAI GPT-6 Astra）は実行しなかった。
- 出力制限を大きくすれば、無回答実行（3.1節）の一部が計画になる可能性がある。

6.4 ケースの出所

すべてのケースは1つのプラットフォームのユーザーからのものである。他の荷主の貨物構成は異なる場合がある。

7. 次のステップ

7.1 実務ルール

エンジンとバリデータv3に、次の2つの実務ルールを追加する：

- **ルーフクリアランス：** ルーフ下に最低50 mm (2.0 in)、または25 kg (55 lb) 超の段積み品の上に100 mm (3.9 in)。
- **サイドプレイ** フォークリフト列ごと。

ルーフルールで全エントラントを再採点した。傾けられない床置き品で、（ルーフの高さ－50 mm (2.0 in)）より高いものは免除される。ケースの制約上やむを得ないためだ。

- **エンジン**は標準ラン16件中9件、大型ラン8件中1件でルールを守っている。
- **6件のランでは、モデルの回答がルールを守っている** エンジンに近い計画で。これらがエンジンの最初のターゲット：

    - H02 (GPT);
    - H06 (Grok, 11.41 m (37.4 ft));
    - H12 LDM (Gemini, 11.00 m (36.1 ft));
    - H12パレット (Kimi, 28ポジション);
    - BV3 LDM (GPT, 9.14 m (30.0 ft));
    - BV5 (Opus, 124個配置)。
- **8件のランでは、いずれのエントラントもルールを守れていない：** H01、H03、H08、BV1、BV2の両ラン、BV3パレット、BV6。

7.2 次版

- ルーフクリアランス、サイドプレイ、積付け順序チェックを備え、全エントラントに適用するバリデータv3。
- 実顧客の新しいケース。
- 各ラボの最も高性能なモデル（プレミアムティア含む）。モデルごとに複数ランし、分散を報告。
- エントラントとしてニューロシンボリックループを追加。

8. 公開

公開:

- 全ケース（完全版）、
- バリデーター、
- プロンプト、
- スコアリング・ランキングコード、
- 各ラン：送信リクエスト、プロバイダーの生レスポンス、解析済みプラン。

ビューアー: 各参加者の全プランを、プロンプトおよび生の回答と並べて3Dで表示：
hansatic.com/ja/neuro-symbolic-load-planning。

非公開:

- パレットポジションカウンター（Hansaticの本番カウンター、すべての参加者に同じ方法で適用）、
- エンジン、
- 元プランのID。

付録A. プロンプトテンプレート

車種ごとにシステムプロンプトを1つ。送信されたものをそのまま掲載。ユーザーメッセージは目的を示し、続けてケースをJSONで示す。品目リストはここでは短縮。プロンプトは英語で表示。モデルが受け取った言語そのまま。

海上コンテナ（BV1、40' ハイキューブ）

システムプロンプト

```
You are an expert load planner. You will be given a sea container and a cargo manifest. Produce a complete,
physically legal loading plan.

Coordinates are millimetres. The origin is the front-left floor corner of the container.
x runs along the length from the front wall (x = 0) to the doors (x = interior length).
y runs across the width from the left wall. z runs up from the floor.
All contact and overlap tests allow 1 mm of tolerance.

Hard rules (a layout that breaks any of them is illegal):
1. Every placed item lies fully inside the cargo space: x ⩾ 0, y ⩾ 0, z ⩾ 0, x + dx ⩽ length, y + dy ⩽
width, z + dz ⩽ height.
2. No two items overlap in 3D.
3. dx, dy, dz are the item's extents along x, y, z after you orient it, and must be a permutation of its length,
width and height. Each item id appears at most once.
4. If tiltable is false the item stays upright: dz must equal its height_mm (it may only turn about the vertical
axis).
5. Support: an item with z > 0 must rest with at least 95% of its base area on the tops of items whose top is
exactly at its z.
6. Nothing may rest on an item with stackable = false, nor on a fragile item that has no max_stack_kg above 0.
7. If an item has max_stack_kg, the total weight of everything at or above its top that overlaps its footprint
(at any height) must not exceed it.
8. Total weight of placed items must not exceed max_payload_kg.
9. Balance. Skipped for a light load: total manifest weight below 15% of max_payload_kg, total manifest volume
below 25% of the cargo space, and no item of 500 kg or more (all three). Otherwise the centre of gravity of
placed cargo plus tare_kg (tare acts at the middle of the floor) must satisfy:
   a. across the width: between 40% and 60% of the width from the left wall;
   b. along the length: between 35% and 65% of the length.
10. Slender piles: an item standing on another item's top, together with the comparable items directly below it,
may be at most 5 times as tall as its narrowest braced width. An item's braced width is the narrower side of the
block it stands in: the items at the same z that touch it side to side (gap at most 10 mm), each at least half as
tall as its neighbour, chained; walls do not count, and a lone item's braced width is its own narrower side. The
pile's narrowest braced width is the smallest over its items. The pile ends at the floor or at an item whose
narrower side is more than 1.5 times the top item's narrower side; a pile that reaches the floor, and a single
item on such a wider base, are not limited.

Loading practice, on the items standing on the floor (reported as a separate verdict, not part of legality):
- They form at most two blocks: a gap longer than 300 mm along x between them starts a new block.
- Items at least 85% of the width wide that share a footprint line up at the same y.

Also reported, never scored: floor load per item against max_floor_load_kg_m2, how far the floor block stands
from the nearest end wall (more than 150 mm), more than 3 gaps over 60 mm across the width in one cross-section,
items that could tip over in transport.

Place as many items as you legally can. Any item you cannot place goes in "unplaced".

Answer with one JSON object and nothing else after it, in this exact shape:
{"placements":[{"id":"I01","x":0,"y":0,"z":0,"dx":1200,"dy":800,"dz":1000}],"unplaced":["I07"]}
```

ユーザーメッセージ（品目リスト短縮）

```
Objective: Minimize loading meters: the occupied length, from the frontmost to the rearmost face of any cargo
along x. Fewer meters is better.
```

```
{
 "vehicle": {
  "code": "40HC",
  "type": "container",
  "interior_mm": {
   "length": 12032,
   "width": 2352,
   "height": 2698
  },
  "max_payload_kg": 26580,
  "tare_kg": 3900,
  "max_floor_load_kg_m2": 5400
 },
 "objective": "ldm",
 "items": [
  {
   "id": "I001",
   "length_mm": 6344,
   "width_mm": 90,
   "height_mm": 638,
   "weight_kg": 31.34,
   "stackable": true,
   "max_stack_kg": 1000,
   "fragile": false,
   "tiltable": true
  },
  {
   "id": "I002",
   "length_mm": 2706,
   "width_mm": 90,
   "height_mm": 354,
   "weight_kg": 12.72,
   "stackable": true,
   "max_stack_kg": 1000,
   "fragile": false,
   "tiltable": true
  },
  "... 101 more items"
 ]
}
```


トラック（H09、13.6 m（44.6 ft）カーテンサイダー）

システムプロンプト

```
You are an expert load planner. You will be given a truck and a cargo manifest. Produce a complete, physically
legal loading plan.

Coordinates are millimetres. The origin is the front-left floor corner of the cargo space.
x runs along the length from the front wall (x = 0) to the doors (x = interior length).
y runs across the width from the left wall. z runs up from the floor.
All contact and overlap tests allow 1 mm of tolerance.

Hard rules (a layout that breaks any of them is illegal):
1. Every placed item lies fully inside the cargo space: x ⩾ 0, y ⩾ 0, z ⩾ 0, x + dx ⩽ length, y + dy ⩽
width, z + dz ⩽ height.
2. No two items overlap in 3D.
3. dx, dy, dz are the item's extents along x, y, z after you orient it, and must be a permutation of its length,
width and height. Each item id appears at most once.
4. If tiltable is false the item stays upright: dz must equal its height_mm (it may only turn about the vertical
axis).
5. Support: an item with z > 0 must rest with at least 95% of its base area on the tops of items whose top is
exactly at its z.
6. Nothing may rest on an item with stackable = false, nor on a fragile item that has no max_stack_kg above 0.
7. If an item has max_stack_kg, the total weight of everything at or above its top that overlaps its footprint
(at any height) must not exceed it.
8. Total weight of placed items must not exceed max_payload_kg.
9. Balance. Skipped for a light load: total manifest weight below 15% of max_payload_kg, total manifest volume
below 25% of the cargo space, and no item of 500 kg or more (all three). Otherwise the centre of gravity of
placed cargo plus tare_kg (tare acts at the middle of the floor) must satisfy:
```

```
   a. across the width: between 40% and 60% of the width from the left wall;
   b. the total weight (cargo + tare) splits between kingpin and bogie like a beam on two supports at
kingpin_from_front_mm and bogie_from_front_mm; neither may exceed its max (not checked when the total is within
both limits).
10. Slender piles: an item standing on another item's top, together with the comparable items directly below it,
may be at most 5 times as tall as its narrowest braced width. An item's braced width is the narrower side of the
block it stands in: the items at the same z that touch it side to side (gap at most 10 mm), each at least half as
tall as its neighbour, chained; walls do not count, and a lone item's braced width is its own narrower side. The
pile's narrowest braced width is the smallest over its items. The pile ends at the floor or at an item whose
narrower side is more than 1.5 times the top item's narrower side; a pile that reaches the floor, and a single
item on such a wider base, are not limited.

Loading practice, on the items standing on the floor (reported as a separate verdict, not part of legality):
- They form one block: a gap longer than 300 mm along x between them starts a new block.
- Items at least 85% of the width wide that share a footprint line up at the same y.

Also reported, never scored: floor load per item against max_floor_load_kg_m2, how far the floor block stands
from the nearest end wall (more than 150 mm), more than 3 gaps over 60 mm across the width in one cross-section,
items that could tip over in transport.

Place as many items as you legally can. Any item you cannot place goes in "unplaced".

Answer with one JSON object and nothing else after it, in this exact shape:
{"placements":[{"id":"I01","x":0,"y":0,"z":0,"dx":1200,"dy":800,"dz":1000}],"unplaced":["I07"]}
```

ユーザーメッセージ（品目リスト短縮）

```
Objective: Minimize loading meters: the occupied length, from the frontmost to the rearmost face of any cargo
along x. Fewer meters is better.

{
 "vehicle": {
  "code": "CURTAINSIDER_136",
  "type": "truck",
  "interior_mm": {
   "length": 13620,
   "width": 2480,
   "height": 2700
  },
  "max_payload_kg": 24000,
  "tare_kg": 7200,
  "max_floor_load_kg_m2": 5000,
  "kingpin_from_front_mm": 840,
  "max_kingpin_kg": 12000,
  "bogie_from_front_mm": 12200,
  "max_bogie_kg": 21000
 },
 "objective": "ldm",
 "items": [
  {
   "id": "I01",
   "length_mm": 1000,
   "width_mm": 600,
   "height_mm": 800,
   "weight_kg": 100,
   "stackable": true,
   "max_stack_kg": 1000,
   "fragile": true,
   "tiltable": false
  },
  {
   "id": "I02",
   "length_mm": 3800,
   "width_mm": 1300,
   "height_mm": 1600,
   "weight_kg": 1000,
   "stackable": true,
   "max_stack_kg": 1000,
   "fragile": false,
   "tiltable": false
  },
  "... 28 more items"
```

```
 ]
}
```

航空コンテナ（U01、斜壁付きAKE）

システムプロンプト

```
You are an expert load planner. You will be given an air container (ULD) and a cargo manifest. Produce a
complete, physically legal loading plan.

Coordinates are millimetres. The origin is the front-left floor corner of the container.
x runs along the length from the front wall (x = 0) to the doors (x = interior length).
y runs across the width from the left wall. z runs up from the floor.
All contact and overlap tests allow 1 mm of tolerance.

Hard rules (a layout that breaks any of them is illegal):
1. Every placed item lies fully inside the cargo space: x ⩾ 0, y ⩾ 0, z ⩾ 0, x + dx ⩽ length, y + dy ⩽
width, z + dz ⩽ height.
   Sloped sides: a sloped side {side, overhang_mm, height_mm} rises from the edge of the flat floor, overhang_mm
in from that wall, to height_mm up at the wall ("y+" is the wall at y = width, "y-" the wall at y = 0). An item
whose bottom is at z may come no closer to that wall than overhang_mm x (1 - z / height_mm); above height_mm the
full width is free.
2. No two items overlap in 3D.
3. dx, dy, dz are the item's extents along x, y, z after you orient it, and must be a permutation of its length,
width and height. Each item id appears at most once.
4. If tiltable is false the item stays upright: dz must equal its height_mm (it may only turn about the vertical
axis).
5. Support: an item with z > 0 must rest with at least 95% of its base area on the tops of items whose top is
exactly at its z.
   Exception for sloped sides: an item may instead lean on the slope with its outer bottom edge, nothing under
it, at z equal to the slope's height under that edge (height_mm x reach / overhang_mm, where reach is how far the
item extends past the edge of the flat floor; within 1.5 mm), provided that z is at most 15% of dy and at least
60% of dy, and the item's centre, lie over the flat floor.
6. Nothing may rest on an item with stackable = false, nor on a fragile item that has no max_stack_kg above 0.
7. If an item has max_stack_kg, the total weight of everything at or above its top that overlaps its footprint
(at any height) must not exceed it.
8. Total weight of placed items must not exceed max_payload_kg.
9. Balance. Skipped for a light load: total manifest weight below 15% of max_payload_kg, total manifest volume
below 25% of the cargo space, and no item of 500 kg or more (all three). Otherwise the centre of gravity of
placed cargo plus tare_kg (tare acts at the middle of the floor) must satisfy:
   a. across the width: between 40% and 60% of the width from the left wall;
   b. along the length: between 35% and 65% of the length.
10. Slender piles: an item standing on another item's top, together with the comparable items directly below it,
may be at most 5 times as tall as its narrowest braced width. An item's braced width is the narrower side of the
block it stands in: the items at the same z that touch it side to side (gap at most 10 mm), each at least half as
tall as its neighbour, chained; walls do not count, and a lone item's braced width is its own narrower side. The
pile's narrowest braced width is the smallest over its items. The pile ends at the floor or at an item whose
narrower side is more than 1.5 times the top item's narrower side; a pile that reaches the floor, and a single
item on such a wider base, are not limited.

Loading practice, on the items standing on the floor (reported as a separate verdict, not part of legality):
- They form at most two blocks: a gap longer than 300 mm along x between them starts a new block.
- Items at least 85% of the width wide that share a footprint line up at the same y.

Also reported, never scored: floor load per item against max_floor_load_kg_m2, how far the floor block stands
from the nearest end wall (more than 150 mm), more than 3 gaps over 60 mm across the width in one cross-section,
items that could tip over in transport.

Place as many items as you legally can. Any item you cannot place goes in "unplaced".

Answer with one JSON object and nothing else after it, in this exact shape:
{"placements":[{"id":"I01","x":0,"y":0,"z":0,"dx":1200,"dy":800,"dz":1000}],"unplaced":["I07"]}
```

ユーザーメッセージ（品目リスト短縮）

```
Objective: Place as many items as you legally can. Only the number of legally placed items is ranked.

{
```

```
 "vehicle": {
  "code": "AKE",
  "type": "uld",
  "interior_mm": {
   "length": 1500,
   "width": 1905,
   "height": 1550
  },
  "max_payload_kg": 1506,
  "tare_kg": 82,
  "max_floor_load_kg_m2": 976,
  "sloped_sides": [
   {
    "side": "y+",
    "overhang_mm": 445,
    "height_mm": 610
   }
  ]
 },
 "objective": "count",
 "items": [
  {
   "id": "I01",
   "length_mm": 500,
   "width_mm": 500,
   "height_mm": 500,
   "weight_kg": 1,
   "stackable": true,
   "max_stack_kg": 1000,
   "fragile": false,
   "tiltable": true
  },
  {
   "id": "I02",
   "length_mm": 500,
   "width_mm": 500,
   "height_mm": 500,
   "weight_kg": 1,
   "stackable": true,
   "max_stack_kg": 1000,
   "fragile": false,
   "tiltable": true
  },
  "... 48 more items"
 ]
}
```

パレット（P01、1200 x 1000 mm（47 x 39 in））

システムプロンプト

```
You are an expert load planner. You will be given a pallet and a cargo manifest. Produce a complete, physically
legal loading plan.

Coordinates are millimetres. The origin is one corner of the pallet deck, on its top surface.
x runs along the deck length, y across the deck width, z up from the top of the deck.
All contact and overlap tests allow 1 mm of tolerance.

Hard rules (a layout that breaks any of them is illegal):
1. Every placed item stays on the deck: cargo may hang past each deck edge by at most 25 mm, so -25 ⩽ x, x + dx
⩽ length + 25, -25 ⩽ y, y + dy ⩽ width + 25. And z ⩾ 0, z + dz ⩽ max_cargo_height_mm.
2. No two items overlap in 3D.
3. dx, dy, dz are the item's extents along x, y, z after you orient it, and must be a permutation of its length,
width and height. Each item id appears at most once.
4. If tiltable is false the item stays upright: dz must equal its height_mm (it may only turn about the vertical
axis).
5. Support: an item with z > 0 must rest with at least 95% of its base area on the tops of items whose top is
exactly at its z.
6. Nothing may rest on an item with stackable = false, nor on a fragile item that has no max_stack_kg above 0.
7. If an item has max_stack_kg, the total weight of everything at or above its top that overlaps its footprint
```

```
(at any height) must not exceed it.
8. Total weight of placed items must not exceed max_payload_kg.
9. Balance. Skipped for a light load: total manifest weight below 15% of max_payload_kg, total manifest volume
below 25% of deck length x deck width x max_cargo_height_mm, and no item of 500 kg or more (all three). Otherwise
the centre of gravity of placed cargo plus tare_kg (tare acts at the middle of the deck) must satisfy:
   a. across the width: between 40% and 60% of the deck width;
   b. along the length: between 35% and 65% of the deck length.
10. Slender piles: an item standing on another item's top, together with the comparable items directly below it,
may be at most 5 times as tall as its narrowest braced width. An item's braced width is the narrower side of the
block it stands in: the items at the same z that touch it side to side (gap at most 10 mm), each at least half as
tall as its neighbour, chained; nothing outside the load counts, and a lone item's braced width is its own
narrower side. The pile's narrowest braced width is the smallest over its items. The pile ends at the deck or at
an item whose narrower side is more than 1.5 times the top item's narrower side; a pile that reaches the deck,
and a single item on such a wider base, are not limited.

Loading practice, on the bottom layer (items at z = 0) (reported as a separate verdict, not part of legality):
- They form one block: a gap longer than 300 mm along x between them starts a new block.
- Items at least 85% of the deck width wide that share a footprint line up at the same y.

Also reported, never scored: how far the bottom layer stands from the nearer deck end (more than 150 mm), more
than 3 gaps over 60 mm across the width in one cross-section, items that could tip over in transport.

Place as many items as you legally can. Any item you cannot place goes in "unplaced".

Answer with one JSON object and nothing else after it, in this exact shape:
{"placements":[{"id":"I01","x":0,"y":0,"z":0,"dx":1200,"dy":800,"dz":1000}],"unplaced":["I07"]}
```

ユーザーメッセージ（品目リスト短縮）

```
Objective: Build a stable interlocked pallet. After the number of placed items, the interlock share is ranked:
the share of items above the bottom layer that each rest on at least two different items directly below
(overlapping each by more than 10 mm in both x and y). Higher is better; a column stack scores 0 %. Ties are
broken by the lower build height.

{
 "vehicle": {
  "code": "IND",
  "type": "pallet",
  "deck_mm": {
   "length": 1200,
   "width": 1000
  },
  "max_cargo_height_mm": 1250,
  "max_payload_kg": 1500,
  "tare_kg": 30,
  "preferred_pattern": "interlocked"
 },
 "objective": "interlock",
 "items": [
  {
   "id": "I01",
   "length_mm": 600,
   "width_mm": 400,
   "height_mm": 120,
   "weight_kg": 25,
   "stackable": true,
   "max_stack_kg": 1000,
   "fragile": false,
   "tiltable": false
  },
  {
   "id": "I02",
   "length_mm": 600,
   "width_mm": 400,
   "height_mm": 120,
   "weight_kg": 25,
   "stackable": true,
   "max_stack_kg": 1000,
   "fragile": false,
   "tiltable": false
  },
  "... 38 more items"
```

```
  ]
}
```

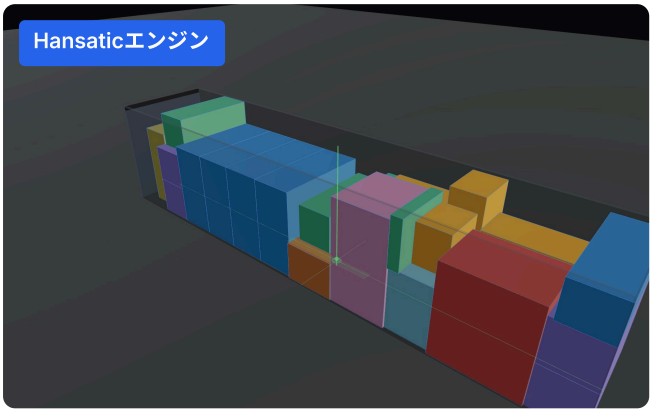
付録B. 全26ラン

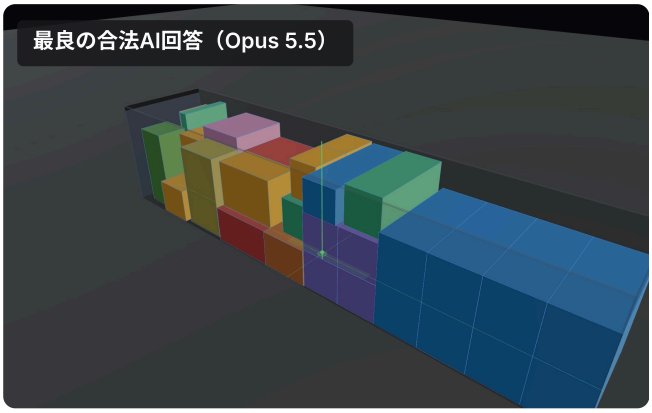
ランごとに1ページ：積荷、エンジンの計画を最良の適法モデル回答と並べ、全参加者の結果を示す。

H01 最短積載長

輸送機器	個数	目的	結果
40' High Cube	27	最短積載長	エンジンがわずかに先行

難しい点
27 個の立て置き必須品





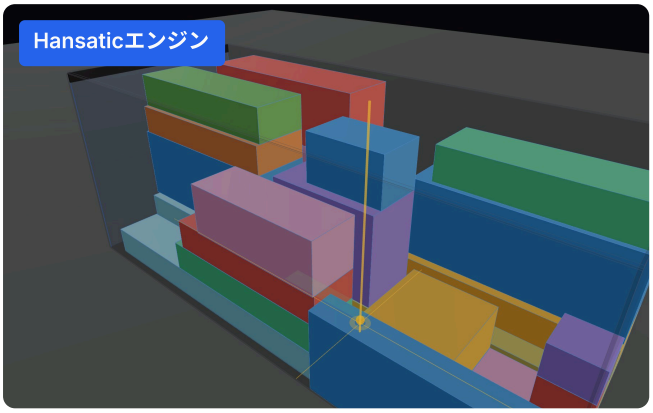
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**26 / 27**	**11.69 m (38.4 ft)**	
GPT-5.5	**合法**	25 / 27	11.88 m (39.0 ft)	
Opus 5.5	**合法**	26 / 27	11.74 m (38.5 ft)	
Grok 4.7	**合法**	25 / 27	11.37 m (37.3 ft)	
Kimi K3.1 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
DeepSeek V4 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
Gemini 4 Argon	**違法**	22 / 27	11.56 m (37.9 ft)	1 貨物サイズの誤り, 1 車両外, 1 宙に浮いた貨物
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	**違法**	27 / 27	12.09 m (39.7 ft)	7 貨物の重複, 14 車両外, 3 高さ制限超過, 1 宙に浮いた貨物, 1 重心ずれ

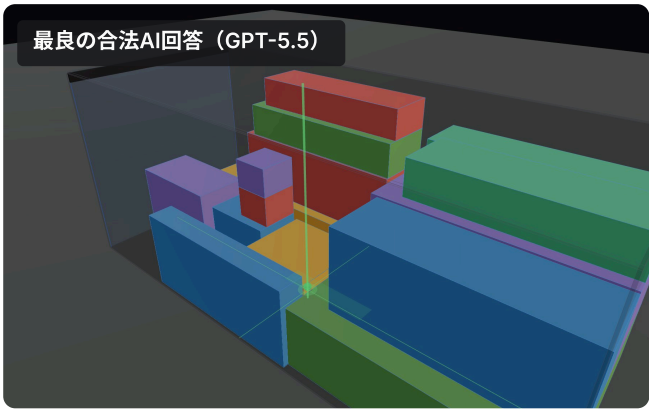
H02 最短積載長

輸送機器	個数	目的	結果
20' Dry Container	29	最短積載長	エンジンが明確に先行

難しい点

29 個の立て置き必須品; 長さ最大5.00 m (16.4 ft)の品目; 最重量品 1,550 kg (3,417 lb)





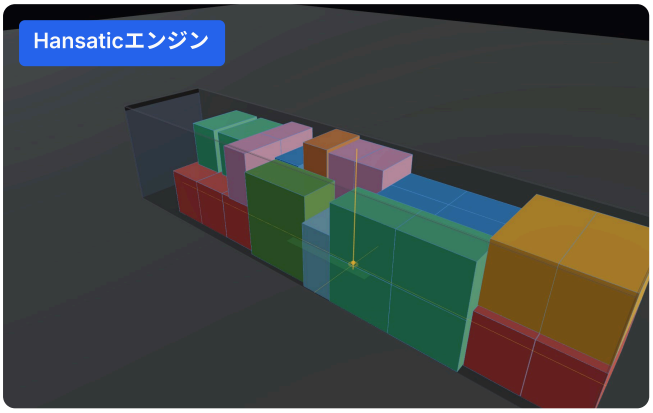
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**24 / 29**	**5.78 m (19.0 ft)**	
GPT-5.5	**合法**	18 / 29	5.76 m (18.9 ft)	
Opus 5.5	計画未返却	-	-	計画を書く前に出力制限に達した。
Grok 4.7	プロバイダーエラー	-	-	プロバイダーがリクエストを拒否した。
Kimi K3.1 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
DeepSeek V4 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
Gemini 4 Argon	**違法**	23 / 29	5.00 m (16.4 ft)	6 貨物サイズの誤り, 13 貨物の重複, 9 宙に浮いた貨物, 9 圧壊貨物
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	**違法**	29 / 29	7.97 m (26.1 ft)	37 貨物の重複, 23 車両外, 1 高さ制限超過, 7 宙に浮いた貨物, 5 圧壊貨物, 2 重心ずれ

H03 最短積載長

輸送機器	個数	目的	結果
40' High Cube	32	最短積載長	エンジンが明確に先行

難しい点

25 個の立て置き必須品; 最重量品 2,680 kg (5,908 lb)



このランでは合法な計画を返したAIモデルはありません。

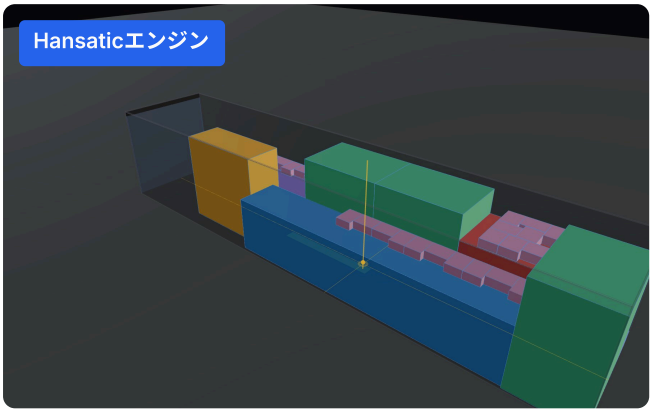
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**31 / 32**	**10.51 m (34.5 ft)**	
GPT-5.5	計画未返却	-	-	計画を書く前に出力制限に達した。
Opus 5.5	計画未返却	-	-	計画を書く前に出力制限に達した。
Grok 4.7	**違法**	31 / 32	12.03 m (39.5 ft)	1 圧壊貨物
Kimi K3.1 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
DeepSeek V4 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
Gemini 4 Argon	**違法**	31 / 32	11.81 m (38.7 ft)	1 貨物サイズの誤り, 12 貨物の重複, 6 車両外, 3 高さ制限超過, 2 宙に浮いた貨物, 3 圧壊貨物
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	計画未返却	-	-	回答に有効な計画が含まれていなかった。

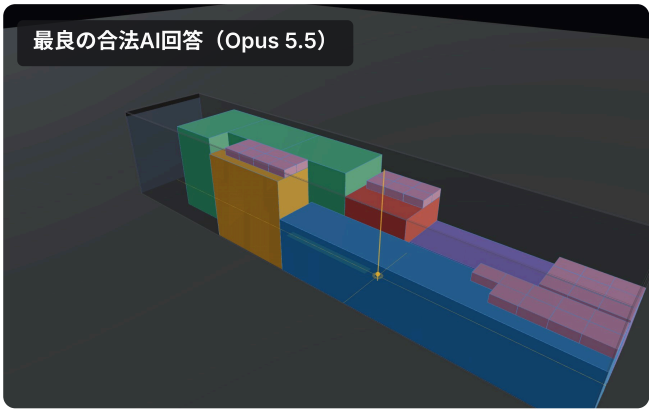
H04 最短積載長

輸送機器	個数	目的	結果
40' Dry Container	39	最短積載長	同点

難しい点

長さ最大6.30 m (20.7 ft)の品目; 最重量品 998 kg (2,200 lb)





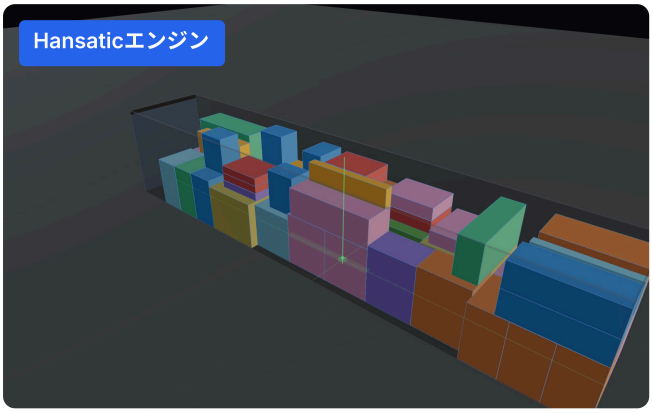
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**39 / 39**	**10.44 m (34.3 ft)**	
GPT-5.5	**合法**	39 / 39	10.72 m (35.2 ft)	
Opus 5.5	**合法**	39 / 39	10.44 m (34.3 ft)	
Grok 4.7	**合法**	39 / 39	11.10 m (36.4 ft)	
Kimi K3.1 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
DeepSeek V4 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
Gemini 4 Argon	**違法**	29 / 39	6.30 m (20.7 ft)	10 貨物サイズの誤り, 6 宙に浮いた貨物
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	**違法**	38 / 39	10.20 m (33.5 ft)	1 貨物サイズの誤り, 27 貨物の重複, 6 車両外, 24 宙に浮いた貨物, 1 圧壊貨物, 1 重心ずれ

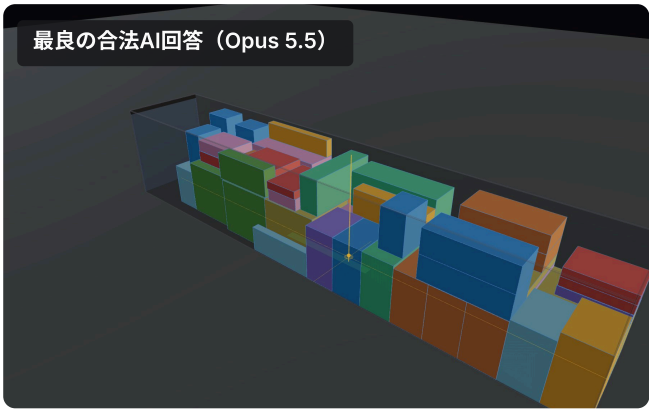
H05 最短積載長

輸送機器	個数	目的	結果
45' High Cube	49	最短積載長	同点

難しい点

49 個の立て置き必須品; 27 個の積み重ね重量制限付き品; 最重量品 814 kg (1,795 lb)





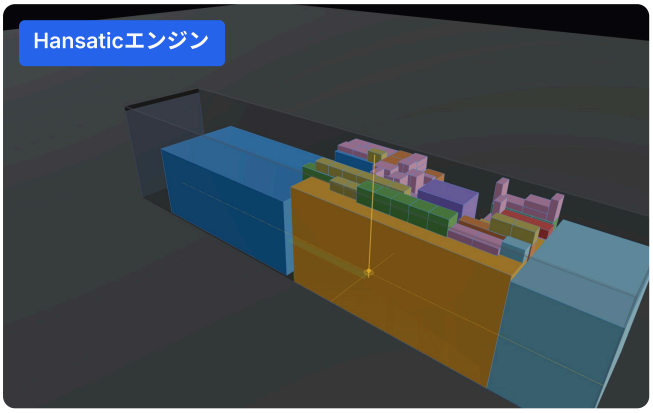
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**49 / 49**	**12.40 m (40.7 ft)**	
GPT-5.5	**合法**	45 / 49	13.28 m (43.6 ft)	
Opus 5.5	**合法**	49 / 49	12.40 m (40.7 ft)	
Grok 4.7	**合法**	43 / 49	12.96 m (42.5 ft)	
Kimi K3.1 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
DeepSeek V4 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
Gemini 4 Argon	**違法**	37 / 49	10.49 m (34.4 ft)	11 貨物サイズの誤り, 2 貨物の重複, 5 車両外, 2 高さ制限超過, 3 宙に浮いた貨物, 2 圧壊貨物
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	**違法**	47 / 49	12.73 m (41.8 ft)	2 貨物サイズの誤り, 31 貨物の重複, 9 車両外, 14 高さ制限超過, 21 宙に浮いた貨物, 2 圧壊貨物

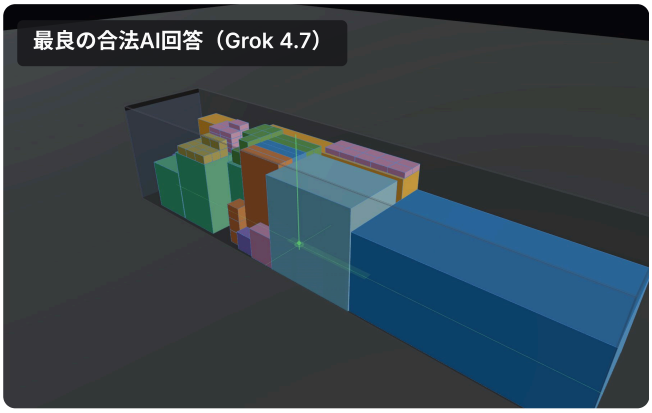
H06 最短積載長

輸送機器	個数	目的	結果
40' High Cube	83	最短積載長	エンジンが明確に先行

難しい点

5 個の積み重ね不可品; 1 個の立て置き必須品; 最重量品 3,020 kg (6,658 lb)





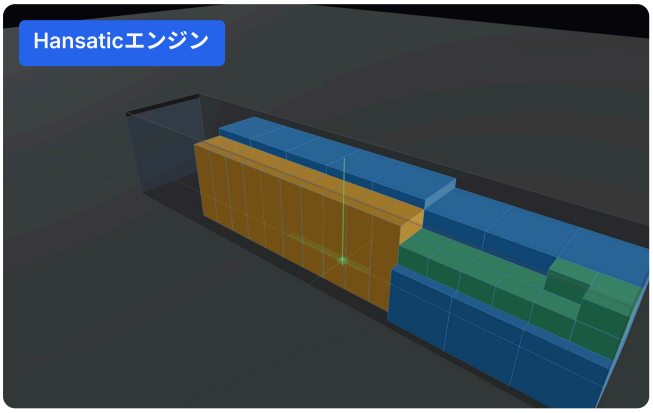
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**83 / 83**	**10.65 m (35.0 ft)**	
GPT-5.5	**合法**	83 / 83	11.78 m (38.7 ft)	
Opus 5.5	計画未返却	-	-	計画を書く前に出力制限に達した。
Grok 4.7	**合法**	83 / 83	11.41 m (37.4 ft)	
Kimi K3.1 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
DeepSeek V4 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
Gemini 4 Argon	**違法**	7 / 83	6.89 m (22.6 ft)	2 貨物サイズの誤り, 1 重心ずれ
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	**違法**	20 / 83	12.25 m (40.2 ft)	2 貨物の重複, 4 車両外, 1 高さ制限超過, 14 宙に浮いた貨物, 1 重心ずれ

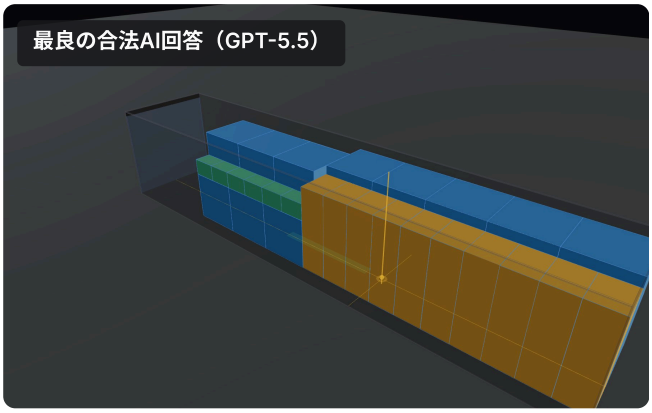
H07 最短積載長

輸送機器	個数	目的	結果
40' Dry Container	45	最短積載長	同点

難しい点

10 個の積み重ね不可品; 15 個の立て置き必須品; 15 個の積み重ね重量制限付き品; 最重量品 1,200 kg (2,646 lb)





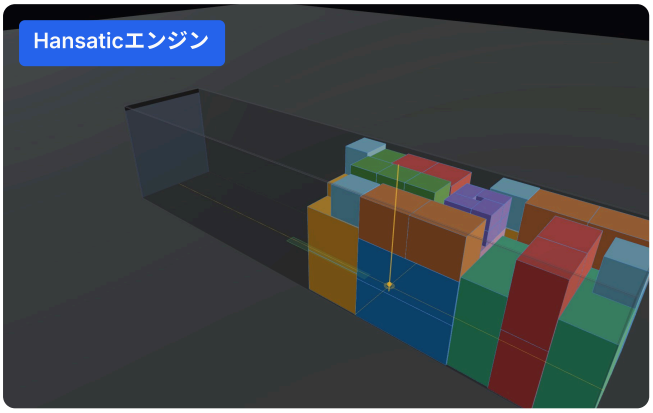
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**45 / 45**	**9.60 m (31.5 ft)**	
GPT-5.5	合法	45 / 45	9.60 m (31.5 ft)	
Opus 5.5	合法	45 / 45	9.61 m (31.5 ft)	
Grok 4.7	違法	45 / 45	10.80 m (35.4 ft)	1 重心ずれ
Kimi K3.1 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
DeepSeek V4 Pro	合法	45 / 45	12.00 m (39.4 ft)	
Gemini 4 Argon	違法	32 / 45	7.60 m (24.9 ft)	13 貨物サイズの誤り, 1 貨物の重複, 16 宙に浮いた貨物
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	違法	10 / 45	12.00 m (39.4 ft)	1 重心ずれ

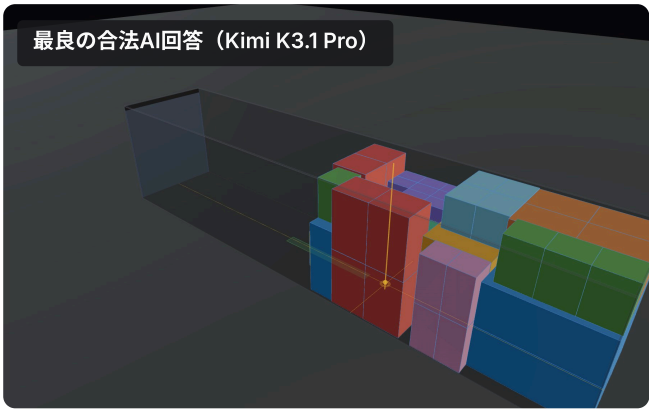
H08 最短積載長

輸送機器	個数	目的	結果
40' High Cube	52	最短積載長	エンジンが明確に先行

難しい点

52 個の立て置き必須品





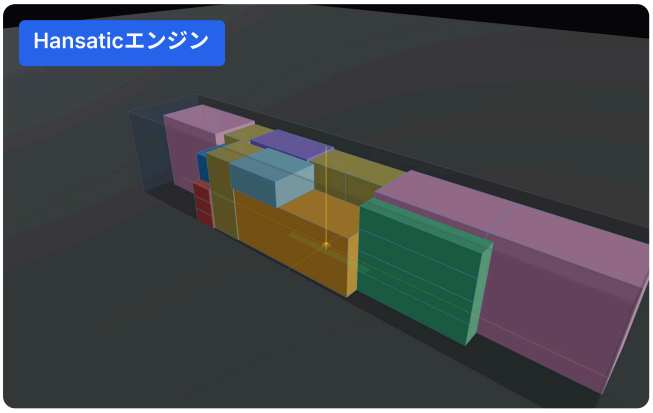
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**52 / 52**	**5.92 m (19.4 ft)**	
GPT-5.5	合法	52 / 52	6.26 m (20.5 ft)	
Opus 5.5	合法	52 / 52	6.20 m (20.3 ft)	
Grok 4.7	合法	52 / 52	6.18 m (20.3 ft)	
Kimi K3.1 Pro	合法	52 / 52	6.12 m (20.1 ft)	
DeepSeek V4 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
Gemini 4 Argon	違法	51 / 52	6.51 m (21.4 ft)	1 貨物サイズの誤り, 4 貨物の重複, 1 宙に浮いた貨物
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	違法	52 / 52	11.25 m (36.9 ft)	14 貨物の重複, 20 車両外

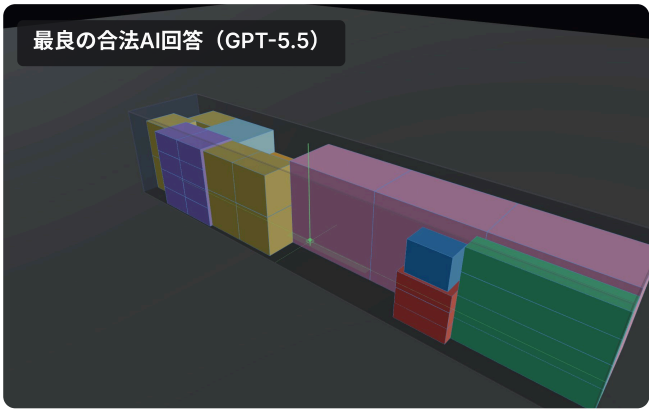
H09 最短積載長

輸送機器	個数	目的	結果
13.6m Curtainsider	30	最短積載長	エンジンがわずかに先行

難しい点

30 個の立て置き必須品; 最重量品 2,400 kg (5,291 lb)





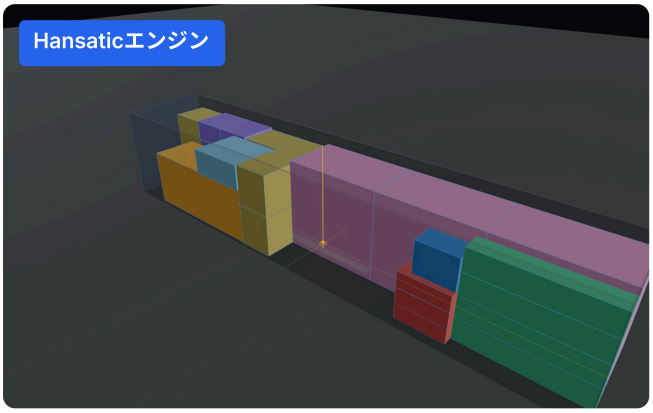
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**29 / 30**	**13.36 m (43.8 ft)**	
GPT-5.5	合法	29 / 30	13.56 m (44.5 ft)	
Opus 5.5	合法	29 / 30	13.60 m (44.6 ft)	
Grok 4.7	合法	29 / 30	13.60 m (44.6 ft)	
Kimi K3.1 Pro	合法	29 / 30	13.60 m (44.6 ft)	
DeepSeek V4 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
Gemini 4 Argon	違法	19 / 30	11.60 m (38.1 ft)	11 貨物サイズの誤り, 7 貨物の重複
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	違法	30 / 30	13.36 m (43.8 ft)	17 貨物の重複, 25 車両外, 7 宙に浮いた貨物, 1 軸重超過

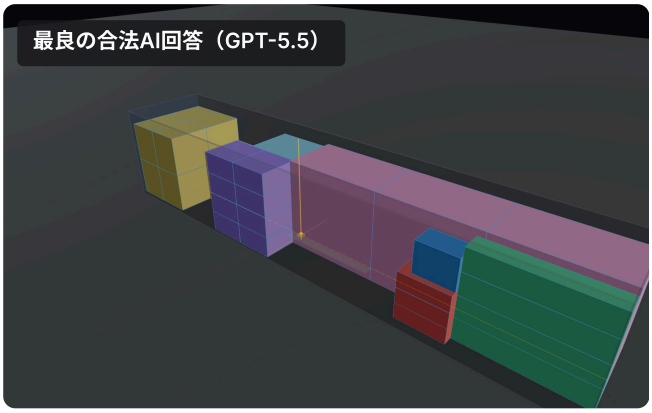
H09 最少パレットポジション

輸送機器	個数	目的	結果
13.6m Curtainsider	30	最少パレットポジション	エンジンがわずかに先行

難しい点

30 個の立て置き必須品; 最重量品 2,400 kg (5,291 lb)





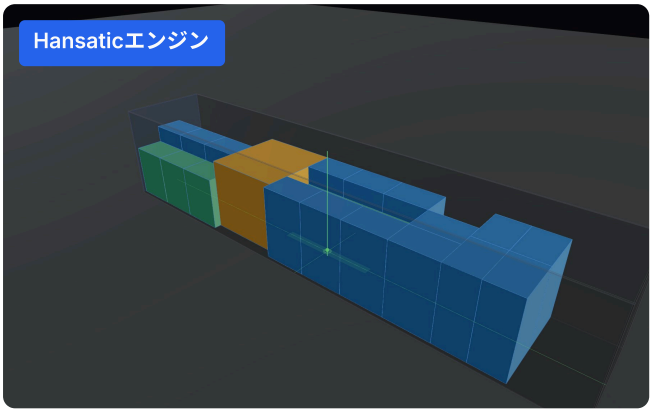
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**29 / 30**	**29 ポジション, 13.36 m (43.8 ft)**	
GPT-5.5	合法	29 / 30	29 ポジション, 13.60 m (44.6 ft)	
Opus 5.5	合法	29 / 30	31 ポジション, 13.60 m (44.6 ft)	
Grok 4.7	合法	29 / 30	29 ポジション, 13.62 m (44.7 ft)	
Kimi K3.1 Pro	合法	29 / 30	29 ポジション, 13.60 m (44.6 ft)	
DeepSeek V4 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
Gemini 4 Argon	違法	30 / 30	29 ポジション, 11.72 m (38.5 ft)	18 貨物の重複, 1 車両外, 3 圧壊貨物, 1 軸重超過
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	違法	30 / 30	34 ポジション, 13.36 m (43.8 ft)	4 貨物の重複, 25 車両外, 1 軸重超過, 1 列のずれ

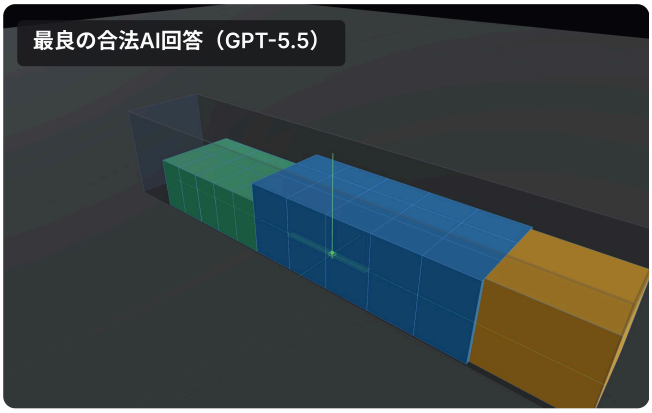
H10 最短積載長

輸送機器	個数	目的	結果
13.6m Curtainsider	26	最短積載長	同点

難しい点

26 個の積み重ね不可品; 16 個の立て置き必須品; 最重量品 666 kg (1,468 lb)





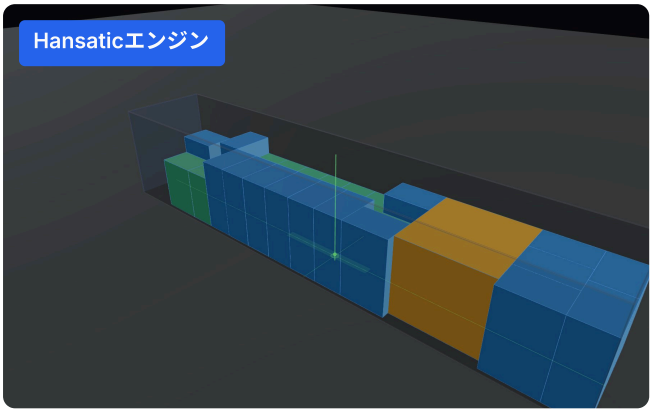
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**26 / 26**	**12.20 m (40.0 ft)**	
GPT-5.5	合法	26 / 26	12.20 m (40.0 ft)	
Opus 5.5	合法	26 / 26	12.20 m (40.0 ft)	
Grok 4.7	合法	26 / 26	12.20 m (40.0 ft)	
Kimi K3.1 Pro	合法	26 / 26	12.20 m (40.0 ft)	
DeepSeek V4 Pro	合法	26 / 26	12.20 m (40.0 ft)	
Gemini 4 Argon	合法	26 / 26	12.60 m (41.3 ft)	
Qwen3p8	合法	26 / 26	12.20 m (40.0 ft)	
Magistral Medium	合法	26 / 26	13.00 m (42.7 ft)	

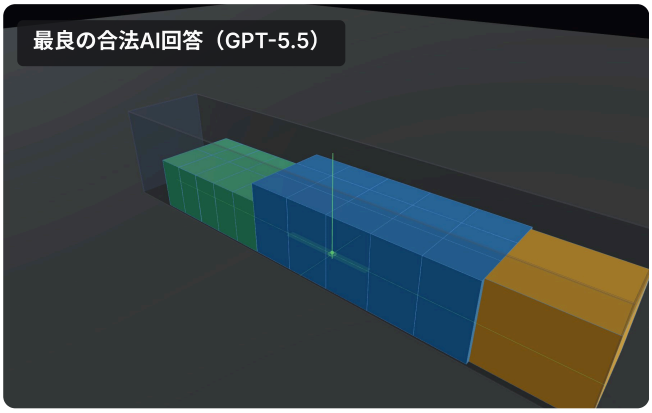
H10 最少パレットポジション

輸送機器	個数	目的	結果
13.6m Curtainsider	26	最少パレットポジション	同点

難しい点

26 個の積み重ね不可品; 16 個の立て置き必須品; 最重量品 666 kg (1,468 lb)





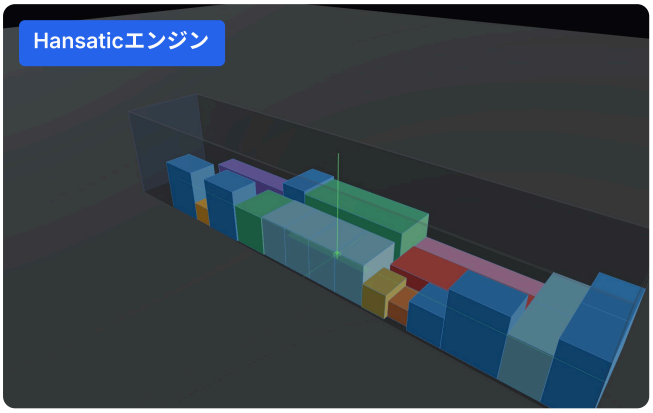
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**26 / 26**	**31 ポジション, 12.20 m (40.0 ft)**	
GPT-5.5	合法	26 / 26	31 ポジション, 12.20 m (40.0 ft)	
Opus 5.5	合法	26 / 26	31 ポジション, 13.00 m (42.7 ft)	
Grok 4.7	合法	26 / 26	31 ポジション, 12.20 m (40.0 ft)	
Kimi K3.1 Pro	合法	26 / 26	31 ポジション, 12.20 m (40.0 ft)	
DeepSeek V4 Pro	合法	26 / 26	31 ポジション, 12.20 m (40.0 ft)	
Gemini 4 Argon	違法	26 / 26	31 ポジション, 12.60 m (41.3 ft)	1 軸重超過
Qwen3p8	合法	26 / 26	31 ポジション, 13.00 m (42.7 ft)	
Magistral Medium	違法	26 / 26	28 ポジション, 13.00 m (42.7 ft)	10 車両外

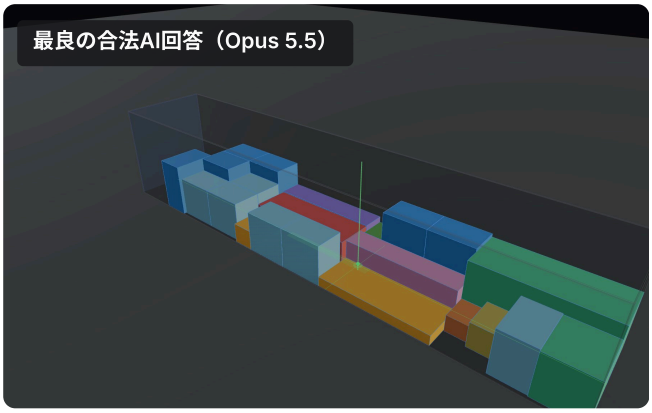
H11 最短積載長

輸送機器	個数	目的	結果
13.6m Curtainsider	23	最短積載長	エンジンがわずかに先行

難しい点

23 個の積み重ね不可品; 23 個の立て置き必須品; 最重量品 2,700 kg (5,952 lb)





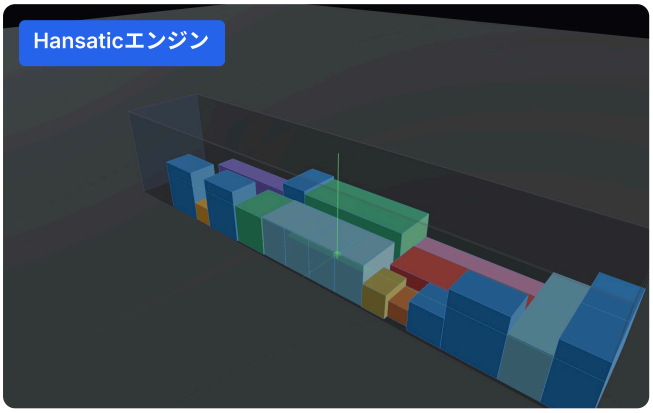
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**23 / 23**	**12.00 m (39.4 ft)**	
GPT-5.5	合法	23 / 23	12.42 m (40.7 ft)	
Opus 5.5	合法	23 / 23	12.20 m (40.0 ft)	
Grok 4.7	合法	23 / 23	12.62 m (41.4 ft)	
Kimi K3.1 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
DeepSeek V4 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
Gemini 4 Argon	合法	23 / 23	13.02 m (42.7 ft)	
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	違法	23 / 23	12.34 m (40.5 ft)	7 貨物の重複, 8 車両外

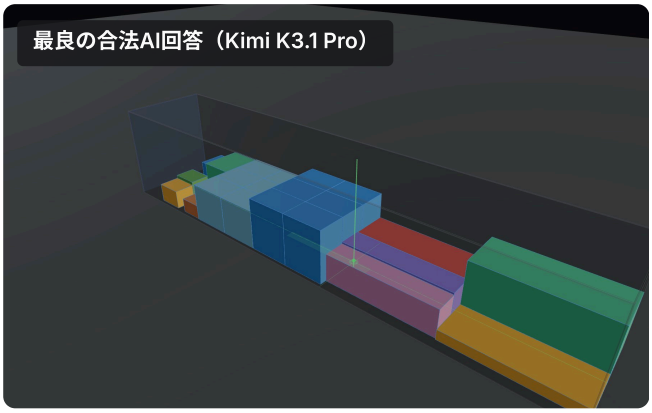
H11 最少パレットポジション

輸送機器	個数	目的	結果
13.6m Curtainsider	23	最少パレットポジション	エンジンが明確に先行

難しい点

23 個の積み重ね不可品; 23 個の立て置き必須品; 最重量品 2,700 kg (5,952 lb)





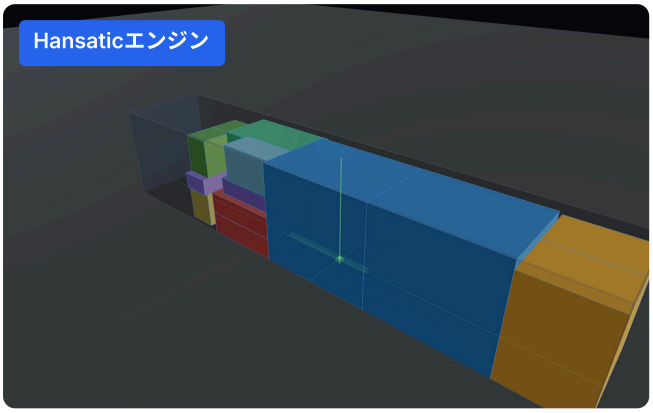
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**23 / 23**	**30 ポジション, 12.00 m (39.4 ft)**	
GPT-5.5	合法	23 / 23	32 ポジション, 12.64 m (41.5 ft)	
Opus 5.5	合法	23 / 23	32 ポジション, 12.84 m (42.1 ft)	
Grok 4.7	プロバイダーエラー	-	-	プロバイダーがリクエストを拒否した。
Kimi K3.1 Pro	合法	23 / 23	32 ポジション, 12.62 m (41.4 ft)	
DeepSeek V4 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
Gemini 4 Argon	違法	22 / 23	30 ポジション, 12.44 m (40.8 ft)	1 貨物サイズの誤り
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	違法	23 / 23	30 ポジション, 12.60 m (41.3 ft)	11 貨物の重複, 9 車両外

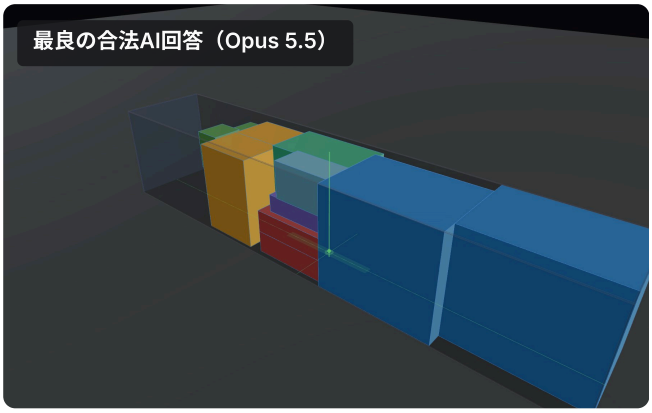
H12 最短積載長

輸送機器	個数	目的	結果
13.6m Curtainsider	14	最短積載長	同点

難しい点

6 個の積み重ね不可品; 9 個の立て置き必須品; 最重量品 732 kg (1,614 lb)





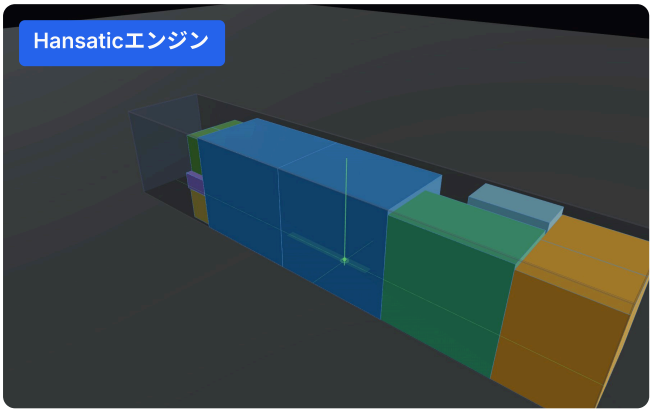
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**13 / 14**	**11.00 m (36.1 ft)**	
GPT-5.5	合法	13 / 14	11.40 m (37.4 ft)	
Opus 5.5	合法	13 / 14	11.00 m (36.1 ft)	
Grok 4.7	プロバイダーエラー	-	-	プロバイダーがリクエストを拒否した。
Kimi K3.1 Pro	合法	13 / 14	11.00 m (36.1 ft)	
DeepSeek V4 Pro	合法	13 / 14	11.70 m (38.4 ft)	
Gemini 4 Argon	合法	13 / 14	11.00 m (36.1 ft)	
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	違法	14 / 14	13.50 m (44.3 ft)	5 貨物の重複, 5 車両外, 2 宙に浮いた貨物

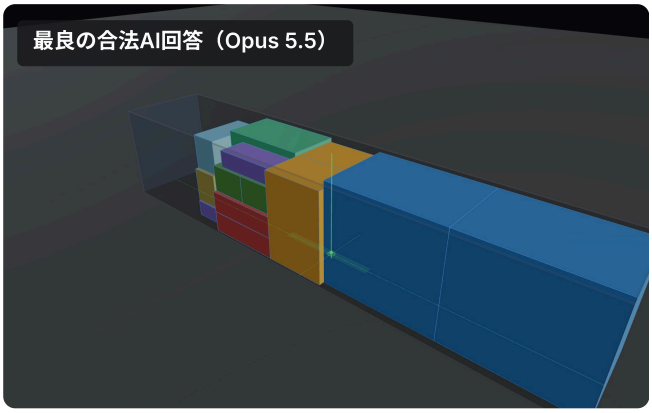
H12 最少パレットポジション

輸送機器	個数	目的	結果
13.6m Curtainsider	14	最少パレットポジション	エンジンが明確に先行

難しい点

6 個の積み重ね不可品; 9 個の立て置き必須品; 最重量品 732 kg (1,614 lb)





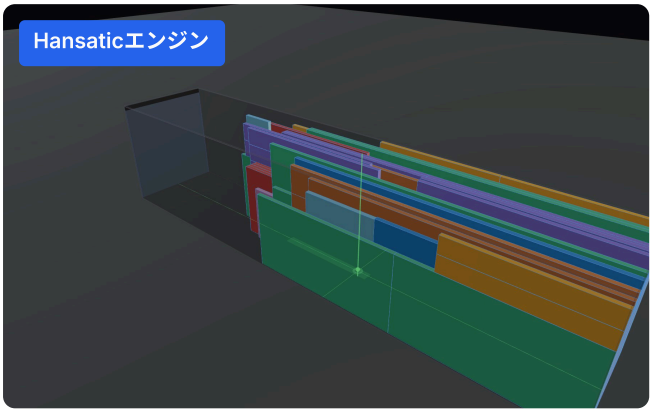
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**13 / 14**	**25 ポジション, 11.00 m (36.1 ft)**	
GPT-5.5	**合法**	13 / 14	29 ポジション, 11.73 m (38.5 ft)	
Opus 5.5	**合法**	13 / 14	28 ポジション, 11.00 m (36.1 ft)	
Grok 4.7	プロバイダーエラー	-	-	プロバイダーがリクエストを拒否した。
Kimi K3.1 Pro	**合法**	13 / 14	28 ポジション, 11.00 m (36.1 ft)	
DeepSeek V4 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
Gemini 4 Argon	**合法**	13 / 14	28 ポジション, 11.15 m (36.6 ft)	
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	**違法**	14 / 14	33 ポジション, 13.20 m (43.3 ft)	9 貨物の重複, 9 車両外

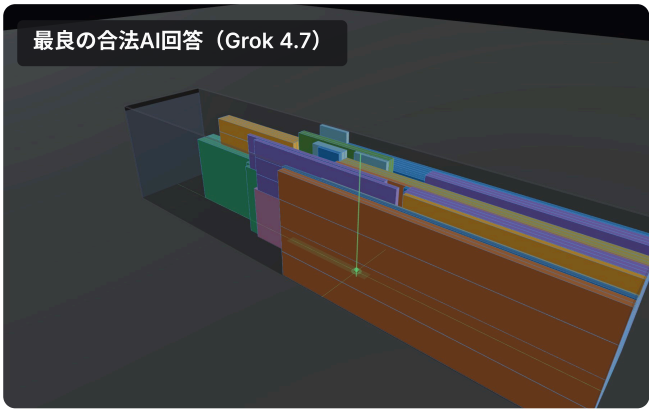
BV1 最短積載長

輸送機器	個数	目的	結果
40' High Cube	103	最短積載長	エンジンが明確に先行

難しい点

37 個の立て置き必須品; 長さ最大8.27 m (27.1 ft)の品目





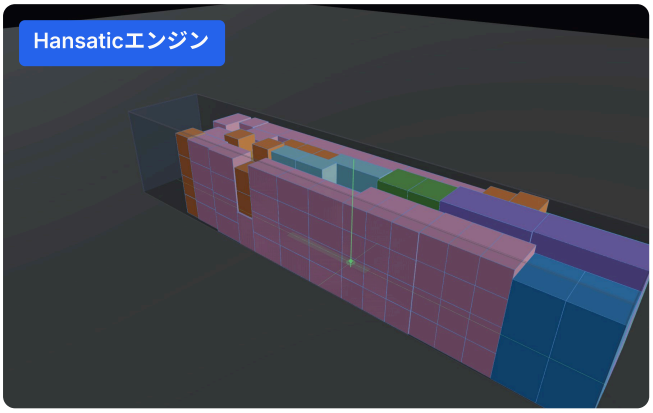
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**103 / 103**	**9.17 m (30.1 ft)**	
GPT-5.5	合法	92 / 103	8.27 m (27.1 ft)	
Opus 5.5	計画未返却	-	-	計画を書く前に出力制限に達した。
Grok 4.7	合法	103 / 103	10.66 m (35.0 ft)	
Kimi K3.1 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
DeepSeek V4 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
Gemini 4 Argon	違法	77 / 103	8.36 m (27.4 ft)	1 貨物サイズの誤り, 4 高さ制限超過, 4 宙に浮いた貨物
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	違法	22 / 103	16.54 m (54.3 ft)	9 車両外

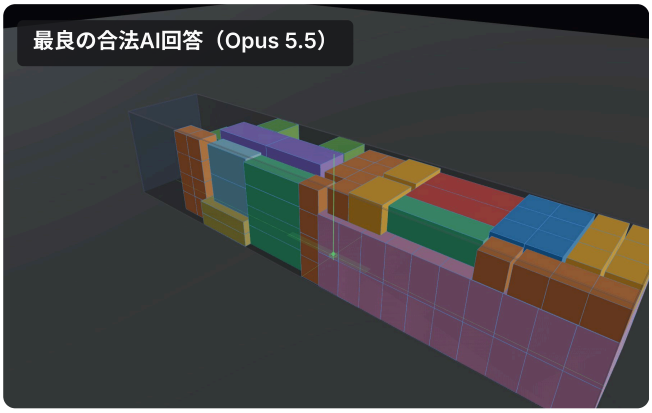
BV2 最短積載長

輸送機器	個数	目的	結果
13.6m Curtainsider	153	最短積載長	エンジンが明確に先行

難しい点

2 個の積み重ね不可品; 最重量品 578 kg (1,274 lb)





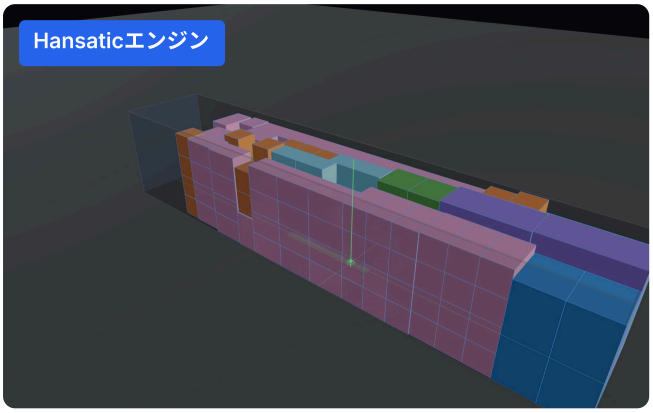
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**153 / 153**	**11.52 m (37.8 ft)**	
GPT-5.5	合法	153 / 153	13.14 m (43.1 ft)	
Opus 5.5	合法	153 / 153	11.78 m (38.6 ft)	
Grok 4.7	合法	153 / 153	12.30 m (40.4 ft)	
Kimi K3.1 Pro	合法	153 / 153	11.99 m (39.3 ft)	
DeepSeek V4 Pro	合法	153 / 153	12.79 m (42.0 ft)	
Gemini 4 Argon	違法	147 / 153	12.77 m (41.9 ft)	6 貨物サイズの誤り
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	違法	56 / 153	14.02 m (46.0 ft)	39 貨物の重複, 27 車両外

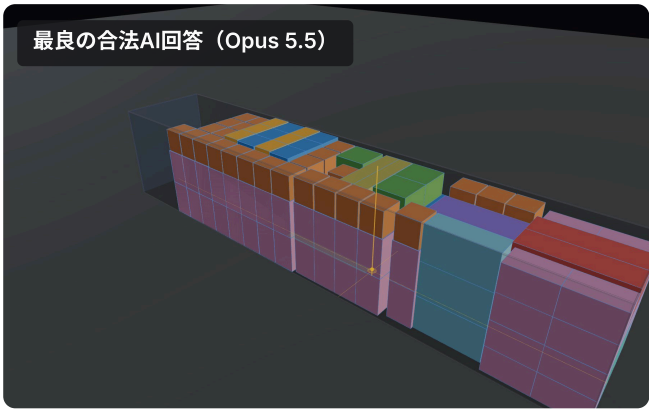
BV2 最少パレットポジション

輸送機器	個数	目的	結果
13.6m Curtainsider	153	最少パレットポジション	エンジンがわずかに先行

難しい点

2 個の積み重ね不可品; 最重量品 578 kg (1,274 lb)





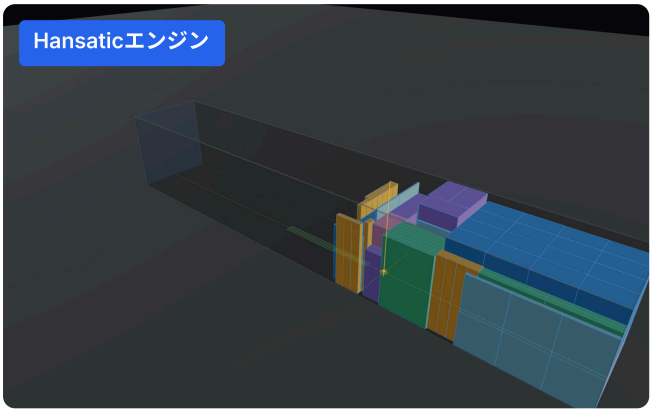
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**153 / 153**	**29 ポジション, 11.52 m (37.8 ft)**	
GPT-5.5	合法	153 / 153	32 ポジション, 12.60 m (41.3 ft)	
Opus 5.5	合法	153 / 153	29 ポジション, 11.67 m (38.3 ft)	
Grok 4.7	合法	153 / 153	32 ポジション, 12.78 m (41.9 ft)	
Kimi K3.1 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
DeepSeek V4 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
Gemini 4 Argon	計画未返却	-	-	計画を書く前に出力制限に達した。
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	違法	153 / 153	34 ポジション, 14.82 m (48.6 ft)	7 貨物の重複, 31 車両外, 86 高さ制限超過, 96 宙に浮いた貨物

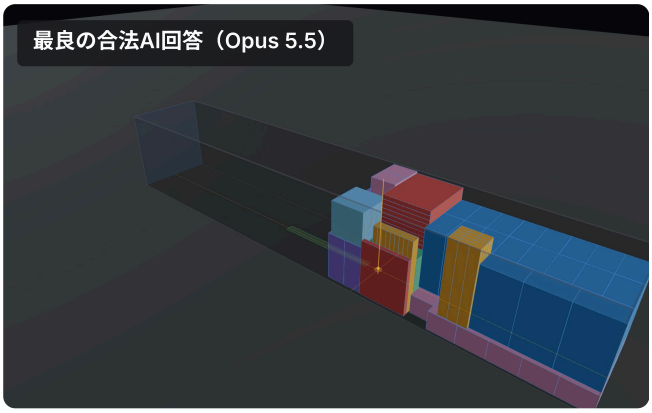
BV3 最短積載長

輸送機器	個数	目的	結果
53ft Dry Van	119	最短積載長	エンジンが明確に先行

難しい点

19 個の積み重ね不可品; 35 個の立て置き必須品





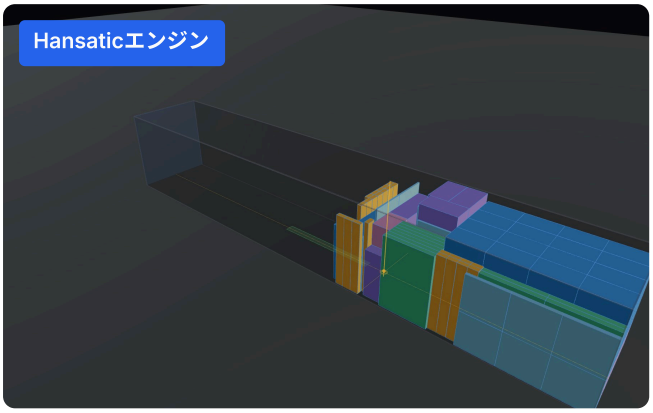
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**119 / 119**	**6.83 m (22.4 ft)**	
GPT-5.5	合法	119 / 119	9.14 m (30.0 ft)	
Opus 5.5	合法	119 / 119	6.99 m (22.9 ft)	
Grok 4.7	合法	119 / 119	8.10 m (26.6 ft)	
Kimi K3.1 Pro	計画未返却	-	-	回答に有効な計画が含まれていなかった。
DeepSeek V4 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
Gemini 4 Argon	違法	115 / 119	8.36 m (27.4 ft)	4 貨物サイズの誤り, 5 貨物の重複, 3 圧壊貨物
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	違法	22 / 119	16.76 m (55.0 ft)	11 車両外

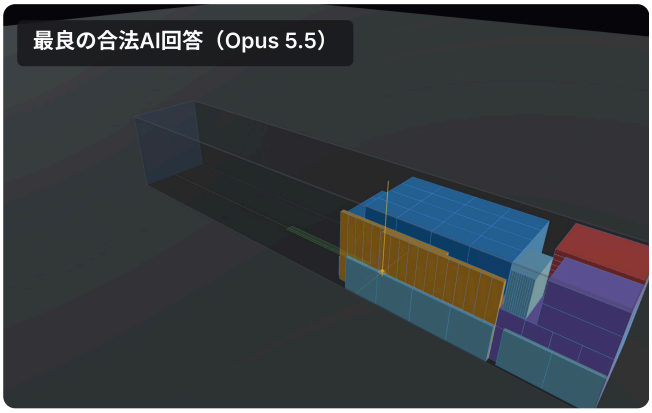
BV3 最少パレットポジション

輸送機器	個数	目的	結果
53ft Dry Van	119	最少パレットポジション	エンジンが明確に先行

難しい点

19 個の積み重ね不可品; 35 個の立て置き必須品





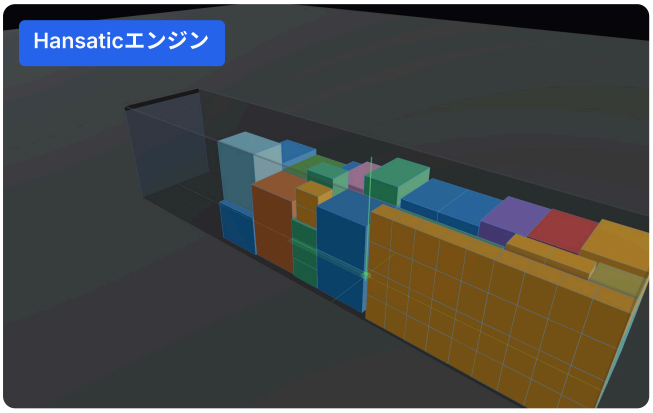
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**119 / 119**	**17 ポジション, 6.83 m (22.4 ft)**	
GPT-5.5	**合法**	119 / 119	23 ポジション, 9.07 m (29.8 ft)	
Opus 5.5	**合法**	119 / 119	18 ポジション, 6.99 m (22.9 ft)	
Grok 4.7	**合法**	119 / 119	22 ポジション, 8.64 m (28.3 ft)	
Kimi K3.1 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
DeepSeek V4 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
Gemini 4 Argon	計画未返却	-	-	計画を書く前に出力制限に達した。
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	**違法**	107 / 119	21 ポジション, 16.18 m (53.1 ft)	12 貨物サイズの誤り, 66 貨物の重複, 35 車両外, 60 宙に浮いた貨物, 1 圧壊貨物, 1 固定されていない緩いグループ

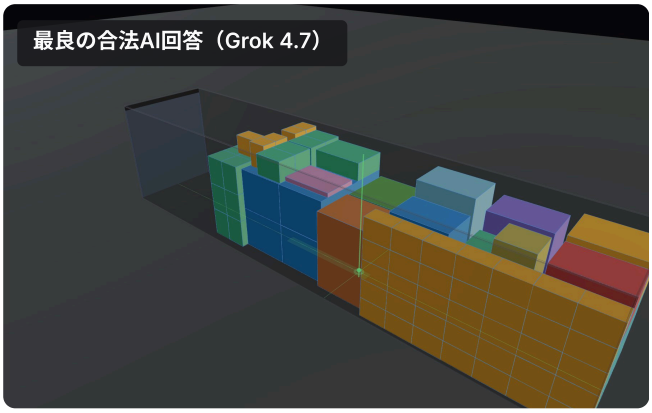
BV4 最短積載長

輸送機器	個数	目的	結果
40' High Cube	145	最短積載長	エンジンが明確に先行

難しい点

11 個の積み重ね不可品; 11 個の壊れ物、最上部のみ; 20 個の立て置き必須品





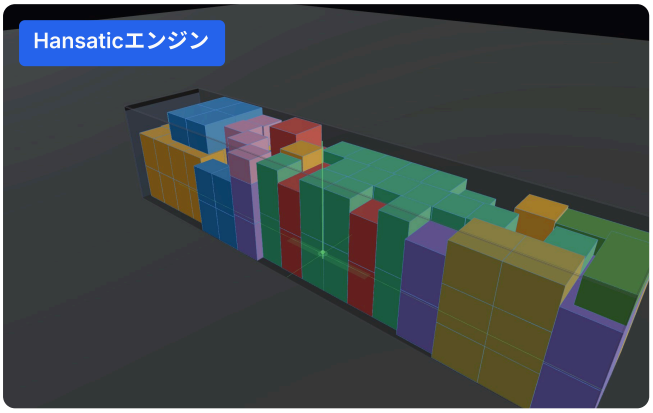
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**145 / 145**	**8.57 m (28.1 ft)**	
GPT-5.5	合法	145 / 145	11.96 m (39.2 ft)	
Opus 5.5	計画未返却	-	-	計画を書く前に出力制限に達した。
Grok 4.7	合法	145 / 145	8.85 m (29.0 ft)	
Kimi K3.1 Pro	合法	145 / 145	10.87 m (35.7 ft)	
DeepSeek V4 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
Gemini 4 Argon	違法	145 / 145	10.05 m (33.0 ft)	19 貨物の重複, 3 車両外, 9 高さ制限超過, 4 宙に浮いた貨物
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	合法	11 / 145	10.33 m (33.9 ft)	

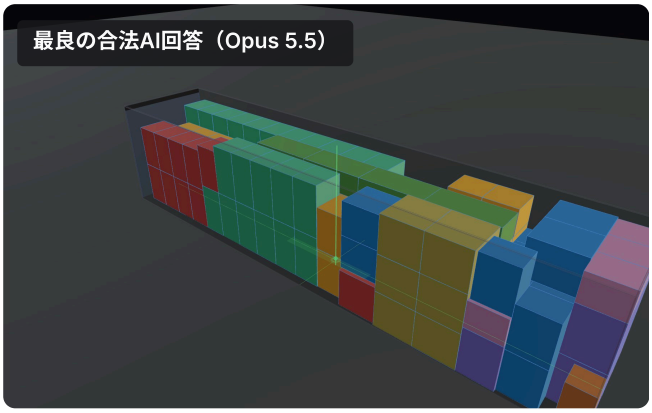
BV5 最短積載長

輸送機器	個数	目的	結果
40' High Cube	136	最短積載長	エンジンが明確に先行

難しい点

136 個の立て置き必須品





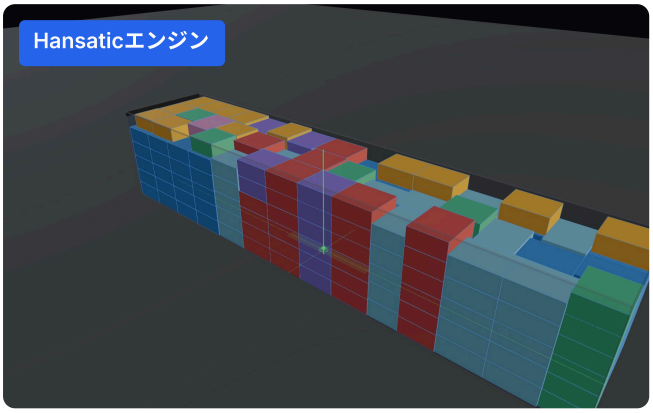
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**127 / 136**	**12.00 m (39.4 ft)**	
GPT-5.5	合法	115 / 136	11.73 m (38.5 ft)	
Opus 5.5	合法	124 / 136	11.77 m (38.6 ft)	
Grok 4.7	合法	115 / 136	11.87 m (39.0 ft)	
Kimi K3.1 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
DeepSeek V4 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
Gemini 4 Argon	違法	121 / 136	12.00 m (39.4 ft)	1 直立必須貨物を傾けた, 12 貨物の重複, 6 車両外, 6 宙に浮いた貨物
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	違法	51 / 136	12.89 m (42.3 ft)	34 車両外, 6 宙に浮いた貨物

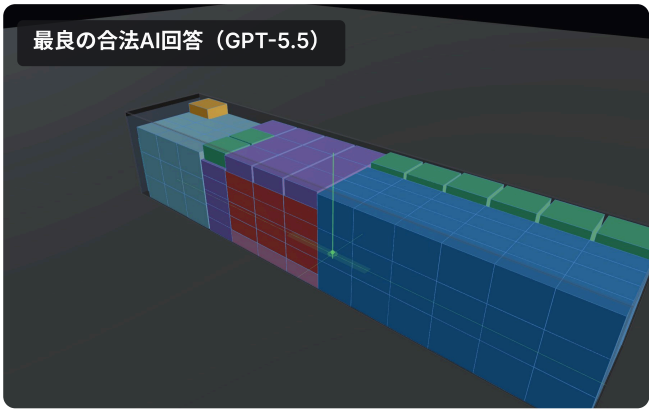
BV6 最短積載長

輸送機器	個数	目的	結果
40' Dry Container	253	最短積載長	エンジンが明確に先行

難しい点

サイズ混在、特別フラグなし



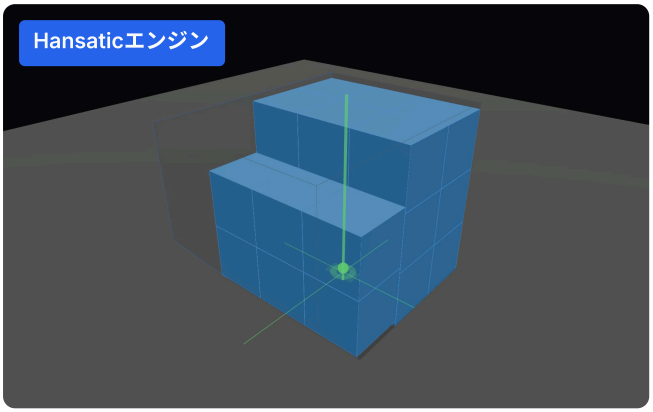


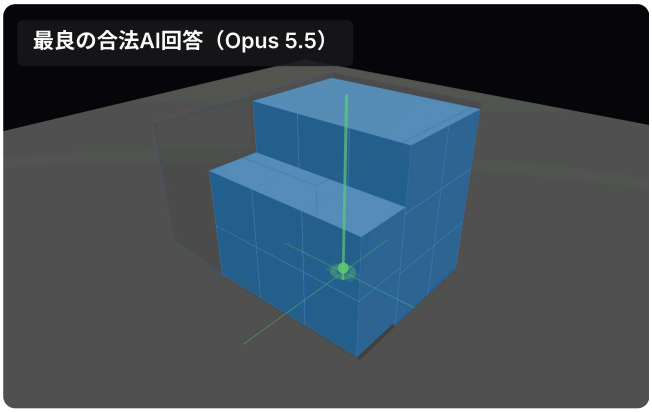
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**244 / 253**	**11.98 m (39.3 ft)**	
GPT-5.5	合法	238 / 253	11.97 m (39.3 ft)	
Opus 5.5	計画未返却	-	-	計画を書く前に出力制限に達した。
Grok 4.7	合法	218 / 253	11.98 m (39.3 ft)	
Kimi K3.1 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
DeepSeek V4 Pro	計画未返却	-	-	計画を書く前に出力制限に達した。
Gemini 4 Argon	違法	197 / 253	11.50 m (37.7 ft)	13 貨物の重複
Qwen3p8	計画未返却	-	-	計画を書く前に出力制限に達した。
Magistral Medium	違法	42 / 253	12.18 m (40.0 ft)	3 車両外

U01 最多配置個数

輸送機器	個数	目的	結果
AKE (LD3)	50	最多配置個数	同点

難しい点
斜壁





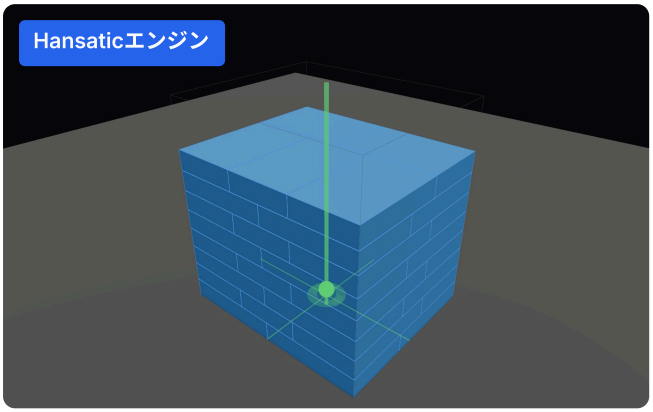
参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**24 / 50**	**24**	
GPT-5.5	合法	18 / 50	18	
Opus 5.5	合法	24 / 50	24	
Grok 4.7	合法	24 / 50	24	
Kimi K3.1 Pro	合法	24 / 50	24	
DeepSeek V4 Pro	合法	18 / 50	18	
Gemini 4 Argon	合法	24 / 50	24	
Qwen3p8	合法	24 / 50	24	
Magistral Medium	違法	24 / 50	24	3 航空コンテナ輪郭外

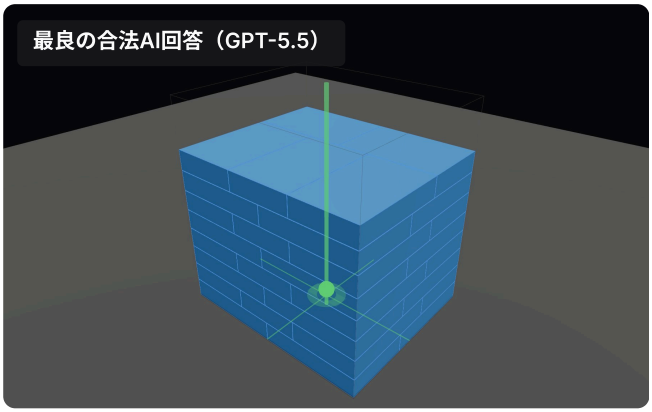
P01 最大インターロックパレット

輸送機器	個数	目的	結果
Industrial / Block 1200×1000	40	最大インターロックパレット	同点

難しい点

40 個の立て置き必須品; 長さ最大0.60 m (2.0 ft)の品目





参加者	ステータス	配置済み	スコア	備考
Hansaticエンジン	**合法**	**40 / 40**	**100%**	
GPT-5.5	合法	40 / 40	100%	
Opus 5.5	合法	40 / 40	100%	
Grok 4.7	合法	40 / 40	100%	
Kimi K3.1 Pro	合法	40 / 40	100%	
DeepSeek V4 Pro	合法	40 / 40	100%	
Gemini 4 Argon	違法	35 / 40	100%	5 貨物サイズの誤り, 12 宙に浮いた貨物
Qwen3p8	合法	40 / 40	100%	
Magistral Medium	違法	40 / 40	100%	30 貨物の重複, 9 車両外