Hansatic

기술 백서

뉴로-심볼릭 적재 계획

실제 고객 화물에 대한 적재 엔진과 8개 프론티어 AI 모델.

2026년 11월

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

모든 계획을 3D로 보여주는 동반 페이지: hansatic.com/ko/neuro-symbolic-load-planning

요약

우리는 결정론적 적재 엔진과 생성형 AI를 결합하고 있습니다. 결합을 만들기 전에 각 구성 요소를 따로 측정했습니다.

- **데이터:** 실제 고객 적재 계획 사례 20건, 채점된 실행 26회. 해상 컨테이너, 커튼사이더, 53 ft (16.2 m) 드라이 밴, AKE 항공 컨테이너, 맞물림 팔레트 적재를 포함합니다.
- **참가자:** Hansatic의 프로덕션 적재 엔진과 8개의 주요 프론티어 LLM. 각 모델은 도구 없이 하나의 JSON 답변으로 모든 품목의 위치와 방향을 제시해야 합니다.
- **심사:** 하나의 결정론적 검증기가 모든 참가자를 동일하게 채점합니다. 다음을 확인합니다:
  - 기하학;
  - 95% 기초 지지;
  - 적재 및 파손 주의 제한;
  - 방향 플래그;
  - 세장한 더미;
  - 화물 중량 및 축하중;
  - 종방향 무게중심 대역.

결과:

- **엔진**은 26회 중 26회에서 적법합니다. 각 실행의 최고 적법 모델 답변과 비교하면 18회에서 우수하고 (13회는 명확히, 5회는 근소하게), 8회는 동점입니다. 뒤처진 적은 없습니다.
- **가장 강력한 모델,** GPT-5.5는 26회 중 25회에서 적법하며, 최대 253개 품목의 화물 명세서를 포함합니다.
- **8개 모델 중 4개**는 실행의 3분의 1 미만에서만 적법합니다.
- **모델의 강점.** 최고 모델 답변은 엔진에 없던 두 가지 배치 전략을 사용했습니다. 둘 다 이제 결정론적 엔진 전략입니다.

우리는 엔진이 오늘날 더 나은 계획자이며 모델의 창의성을 포착할 가치가 있다고 결론짓습니다. 이를 위한 아키텍처는 뉴로-심볼릭입니다:

- 모델이 전략을 제안하고;
- 엔진이 모든 계획을 작성하며;
- 검증기가 적법한 개선만 유지합니다.

사례, 검증기, 프롬프트, 채점 코드를 공개합니다.

1. 벤치마크 설계

1.1 케이스

모든 케이스는 Hansatic 앱의 실제 고객 적재물이며, 1:1로 복사했습니다: 치수, 무게, 수량, 플래그, 차량. 품목 ID는 교체하고 화물명은 제거했습니다.

세트	케이스	런	차량	개품
표준 (H01-H12)	12	16	20' 1대, 40' 드라이 2대, 40' HC 4대, 45' HC 1대, 13.6 m (44.6 ft) 커튼사이더 4대	14-83
대형 (BV1-BV6)	6	8	40' HC 3대, 40' 드라이 1대, 13.6 m (44.6 ft) 커튼사이더 1대, 53 ft (16.2 m) 드라이 밴 1대	103-253
특수 (U01, P01)	2	2	경사면이 있는 AKE (LD3); 1200 × 1000 mm (47 × 39 in) 파렛	50 및 40

런과 목표:

- **트럭**은 두 가지로 채점합니다: 적재 미터(LDM)와 파렛 위치. 표준 케이스 12개에서 16런, 대형 케이스 6개에서 8런이 나옵니다.
- **BV5와 BV6**는 전체가 들어가지 않습니다. 거기서는 더 많은 품목 배치가 주요 목표입니다.
- **U01과 P01**은 각각 알려진 최적값이 있습니다:
  - U01: 경사면 때문에 더 이상 들어갈 수 없어 50개 중 24개;
  - P01: 100% 인터락으로 백 40개 중 40개.

케이스의 제약 조건:

- 중첩 불가 및 상단 전용(파손 주의) 품목;
- 품목별 적재 중량 제한;
- 상향 및 기울기 금지 플래그;
- 긴 바(BV1, 컨테이너 길이까지);
- 단일 고중량 품목(H02);
- 혼합 파렛 및 카톤 적재물(BV3).

1.2 검증기

검증기(v2.3.0)는 모든 계획을 동일하게 검사합니다. 모든 하드 체크를 통과해야만 계획이 적법합니다.

검사 항목	규칙
경계	AKE 경사면을 포함해 모든 품목이 화물 공간 안에 있음
겹침	두 품목이 부피를 공유하지 않음

검사 항목	규칙
지지	각 품목 바닥의 95% 이상이 바닥 또는 아래 품목 위에 놓임
중첩	중첩 불가 품목은 아무것도 올리지 않음; 상단 전용 품목은 아무것도 올리지 않고 맨 위에 놓임; 품목별 적재 중량 제한 준수
방향	상향 및 기울기 금지 플래그 준수
세장한 더미	자립형 적재 더미의 높이-바닥 비율 제한
적재 중량	총 질량이 차량 제한 이내
축 (트럭)	품목 무게중심으로 계산한 전륜·후륜 축하중이 제한 이내
CoG (컨테이너)	종방향 무게중심이 허용 구간 이내
다중 정차	정차지가 있는 경우 정차지별 후입선출

전복 위험은 계산되어 보고되지만 채점하지 않습니다.

1.3 순위

런별 적용:

1. 적법한 계획이 부적법한 계획을 이깁니다.
2. 더 많은 품목을 배치하면 이깁니다.
3. 런의 목표가 정합니다:
    - 더 짧은 적재(LDM);
    - 또는 더 적은 파렛 위치;
    - 또는 P01는 더 높은 인터락 비율, 그다음 더 낮은 높이.
4. 파렛 위치 동점은 LDM으로 깨집니다.

모든 참가자의 파렛 위치를 세는 카운터 하나: Hansatic의 프로덕션 카운터.

마진. 이김의 기준:

- **명확**: LDM 2% 이상, 파렛 위치 1개 이상, 또는 배치 품목 1개 이상;
- **근소**: 그 외.

엔진은 시간 예산 안에서 실행되므로, 단일 런에서 2% 미만 차이는 런 간 편차일 수 있습니다(6.2절).

1.4 참가자

모델. 8개 랩의 주요 프런티어 모델:

- 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)

설정:

- **런:** 케이스당 1회.
- **출력 제한:** 64k 토큰 또는 제공자의 더 낮은 최대값(Gemini: 32k).
- **배치:** 제공되는 경우 제공자 배치 API.
- **답변 수집:** 2026-11-11.
- **총 비용:** 약 $64.

엔진. Hansatic 프로덕션 패커, 2026-11-20 버전(commit 9f66537c). 설정:

- 프로덕션 구성: 워커 4개, 프로덕션 시간 예산, 단일 CPU 코어 고정;
- 케이스당 1런, 베스트-오브 없음;
- 자체 60% 지지 규칙으로 패킹하며, 다른 모든 모델처럼 95%로 심사받음.

1.5 프롬프트

- 차량 종류별 프롬프트 하나: 컨테이너, 트럭, ULD, 파렛.
- 각 프롬프트는 해당 종류에 검증기가 적용하는 정확한 규칙, 목표, 출력 스키마(품목 ID, 위치, 방향)를 명시합니다.
- 해결된 배치 예시는 없습니다.

전체 템플릿은 부록 A에 있습니다.

2. 결과

2.1 모델별, 전체 26회 실행

모델	적법	엔진 동점	엔진 뒤처짐
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

어떤 모델도 엔진을 이기지 못했습니다.

표준 세트 (16회).

모델	적법	동점	뒤처짐
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

대형 세트 (8회). 대형 실행에서 엔진과 동점인 모델은 없습니다.

모델	적법
GPT-5.5	8
Grok 4.7	8
Opus 5.5	5

모델	적법
Kimi K3.1 Pro	2
DeepSeek V4 Pro	1
Magistral Medium	1
Gemini 4 Argon	0
Qwen3p8	0

특수 세트 (2회).

모델	적법	최적 도달
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 엔진 대 최고 적법 모델 답안

세트	명확히 앞섬	근소 앞섬	동점	뒤처짐
표준 (16)	7	3	6	0
대형 (8)	6	2	0	0
특수 (2)	0	0	2	0
전체 (26)	**13**	**5**	**8**	**0**

표준 세트. 각 셀은 배치 품목 수와 적재 길이 또는 팔레트 위치를 보여줍니다.

실행	엔진	최고 적법 모델	차이
H01	26/27, 11.69 m (38.4 ft)	Opus 26, 11.74 m (38.5 ft)	근소 (0.4%)
H02	24/29, 5.78 m (19.0 ft)	GPT 18, 5.76 m (18.9 ft)	명확 (+6개 배치)
H03	31/32, 10.51 m (34.5 ft)	적법 없음	명확
H04	39, 10.44 m (34.3 ft)	Opus 39, 10.44 m (34.3 ft)	동점

실행	엔진	최고 적법 모델	차이
H05	49, 12.40 m (40.7 ft)	Opus 49, 12.40 m (40.7 ft)	동점
H06	83, 10.65 m (35.0 ft)	Grok 83, 11.41 m (37.4 ft)	명확 (6.6%)
H07	45, 9.60 m (31.5 ft)	GPT 45, 9.60 m (31.5 ft)	동점
H08	52, 5.92 m (19.4 ft)	Kimi 52, 6.12 m (20.1 ft)	명확 (3.3%)
H09 LDM	29/30, 13.36 m (43.8 ft)	GPT 29, 13.56 m (44.5 ft)	근소 (1.5%)
H09 팔레트	29/30, 29 위치	GPT 29, 29 위치	근소 (LDM 동점 결승)
H10 LDM	26, 12.20 m (40.0 ft)	DeepSeek, 동일	동점
H10 팔레트	26, 31 위치	DeepSeek, 동일	동점
H11 LDM	23, 12.00 m (39.4 ft)	Opus 23, 12.20 m (40.0 ft)	근소 (1.6%)
H11 팔레트	23, 30 위치	Kimi 23, 32 위치	명확 (2 위치)
H12 LDM	13/14, 11.00 m (36.1 ft)	Gemini 13, 11.00 m (36.1 ft)	동점
H12 팔레트	13/14, 25 위치	Kimi 13, 28 위치	명확 (3 위치)

대형 세트.

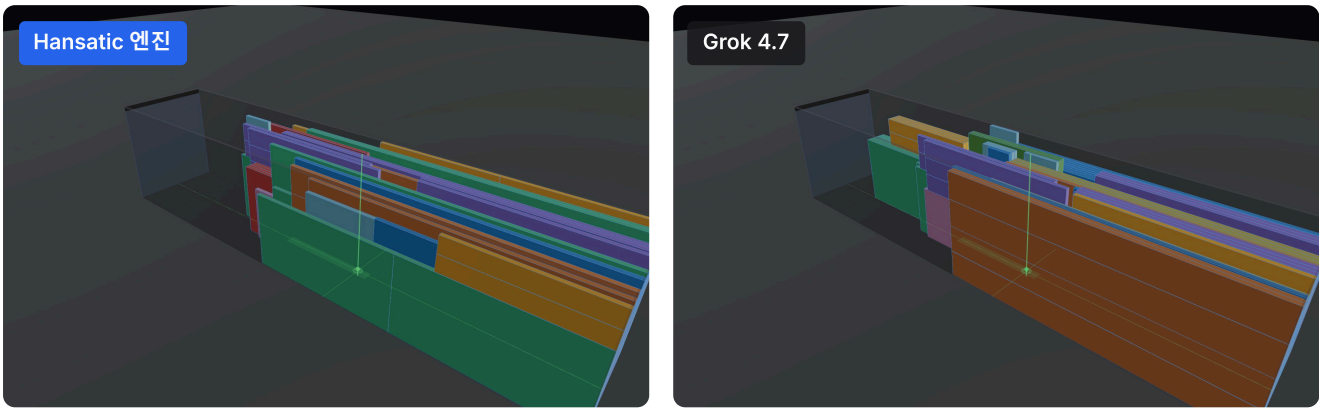
실행	엔진	최고 적법 모델	차이
BV1 (40' HC, 103 bars)	103, 9.17 m (30.1 ft)	Grok 103, 10.66 m (35.0 ft)	명확 (14.0%)
BV2 LDM (커튼사이더, 153)	153, 11.52 m (37.8 ft)	Opus 153, 11.78 m (38.6 ft)	명확 (2.2%)
BV2 팔레트	153, 29 위치	Opus 153, 29 위치	근소 (LDM 동점 결승)
BV3 LDM (53 ft 밴, 119)	119, 6.83 m (22.4 ft)	Opus 119, 6.99 m (22.9 ft)	명확 (2.2%)
BV3 팔레트	119, 17 위치	Opus 119, 18 위치	명확 (1 위치)
BV4 (145, 상단 전용 취급주의)	145, 8.57 m (28.1 ft)	Grok 145, 8.85 m (29.0 ft)	명확 (3.1%)
BV5 (136, 수직 전용)	127, 12.00 m (39.4 ft)	Opus 124, 11.77 m (38.6 ft)	명확 (+3개 배치)
BV6 (253, 부피의 95%)	244, 11.98 m (39.3 ft)	GPT 238, 11.97 m (39.3 ft)	명확 (+6개 배치)

특수 세트.

실행	엔진	모델
U01 (AKE)	24/50, 최적값	Opus, Grok, Kimi, Qwen, Gemini도 24 도달; GPT와 DeepSeek는 18 도달
P01 (팔레트)	40/40, 100% 인터락, 최적값	모든 적법 모델 답안이 최적값에서 40/40 도달

2.3 실제 화물에서 이 격차의 의미

- **BV1:** 1.49 m (4.9 ft) 더 짧음, 40' 컨테이너 길이의 약 12%.
- **H12 팔레트:** 팔레트 위치 3개 적음, 13.6 m (44.6 ft) 커튼사이더 데크의 약 9%.
- **BV6:** 전체가 들어가지 않는 화물에서 6개 더 배치, 따라서 다른 차량에 실 화물이 6개 줄어듦.



BV1: Hansatic 엔진 9.17 m (30.1 ft), 왼쪽; 최고 적법 모델 답안 (Grok 4.7) 10.66 m (35.0 ft), 오른쪽.

2.4 비용과 시간

	엔진	모델
계획당 시간	무시 가능	비교 안 함: 배치 API가 요청을 대기열에 쌓음
계획당 비용	무시 가능	$0.10-0.80

3. 실패 분석

208회 모델 실행(8개 모델 x 26회)에서 95개 계획이 적법합니다. 나머지 113개는 세 가지 다른 방식으로 실패합니다.

모델	적법	답변 없음 (출력 한도)	제공자 오류	규칙 위반 계획
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
합계	**95**	**65**	**4**	**44**

3.1 답변 없음: 모델이 생각할 공간을 다 써버림

65회 실행에서 모델은 출력 한도에 도달할 때까지 추론했습니다. 계획은 작성하지 않았습니다. 한 건 더 (Magistral, H03)는 정상 종료했지만 유효한 JSON이 아니었습니다. 여기에 포함했습니다.

- 가장 흔한 실패입니다. 8개 모델 중 6개에 해당합니다.
- Kimi, DeepSeek, Qwen은 대부분의 실행을 이렇게 잃습니다. 적법한 계획은 대부분 작은 건에서 나옵니다.
- Opus는 이렇게 6회를 잃습니다. 그중 BV1, BV4, BV6은 가장 큰 화물 목록 세 개입니다.

생산 환경의 플래너에게 답변이 없는 것은 부적법한 것과 같은 실패입니다. 하역장에는 여전히 계획이 필요합니다.

3.2 제공자 오류

Grok은 4회 실행에서 HTTP 403을 반환했습니다. 제공자 지출 한도에 도달했기 때문입니다. 이 실행들은 재전송되지 않았고 실패로 채점됩니다.

3.3 규칙을 어기는 계획

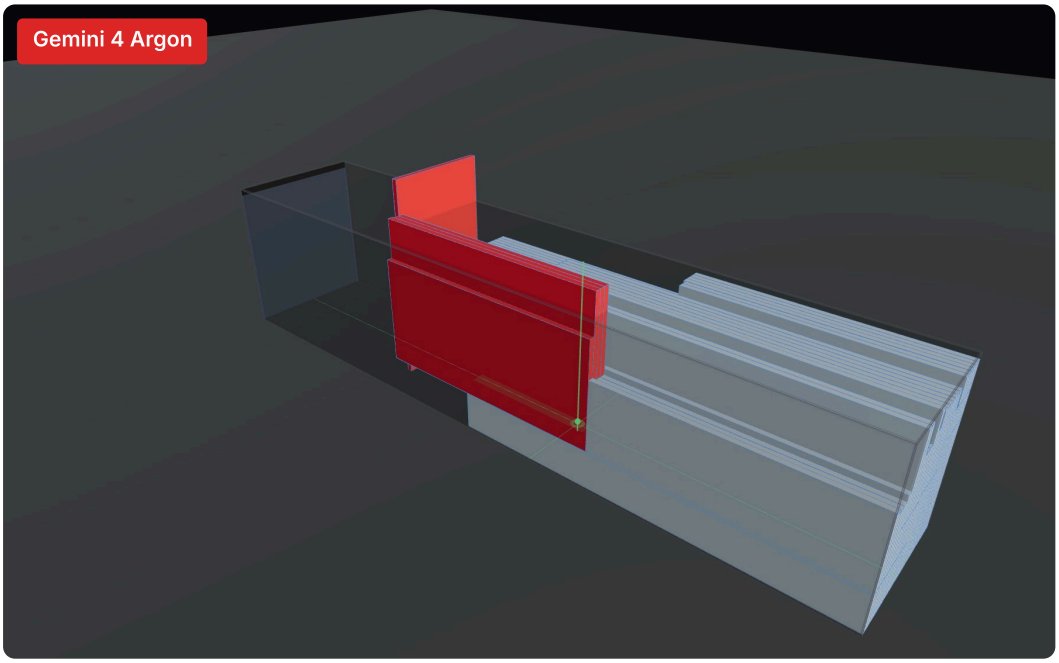
44개 답변은 하나 이상의 엄격한 규칙을 어기는 완전한 계획입니다. 거의 대부분 두 모델에서 나옵니다: Gemini(19), Magistral(23). Grok은 2개입니다.

이 44개 계획 중 각 종류별로 하나 이상 어긴 계획 수:

어긴 규칙	계획 수
적재 공간 밖의 화물	27
겹치는 화물	27
떠 있는 화물 (지지율 95% 미만)	21
허용 방향에 맞지 않는 화물 치수	17
비적재형 또는 취급주의 화물 위 적재, 또는 적재 한도 초과	10
지붕 위 적재	9
무게중심이 허용 범위 밖	7
축하중 초과	4
기울임 금지 화물이 기울어짐	1
AKE 윤곽 밖	1

Grok의 부적법한 두 계획은 각각 하나의 규칙을 어깁니다. 한 실행에서는 화물이 눌리고, 다른 실행에서는 무게중심입니다.

이 계획들은 3D 보기에서 그럴듯해 보입니다. 결함은 종종 소수 화물에 국한됩니다. 수 센티미터 떠 있거나 겹칩니다. 전체 기하학적 검사만이 이를 찾아냅니다.



BV1, Gemini 4 Argon: 103개 화물 중 77개 배치. 강조 표시: 떠 있거나, 겹치거나, 적재 공간을 벗어난 화물.

3.4 적법하지만 뒤처짐

- **불완전함.** BV5와 BV6에서 최고 적법 모델 답변은 엔진보다 3개, 6개 적은 화물을 배치합니다.
- **더 김.** 모든 화물이 들어가는 실행에서 최고 적법 모델 답변은 24개 트럭 및 컨테이너 실행 중 13개에서 엔진보다 깁니다. 가장 큰 차이:

- BV1에서 1.49 m (4.9 ft);
- H06에서 0.76 m (2.5 ft).

3.5 모델이 잘하는 것

가장 뛰어난 답변들은 숙련된 작업반의 전략을 사용합니다:

- **무거운 상자를 앞벽에 붙이기.**
- **비적재형 화물 위 공간 활용.** BV3에서 Opus는 카톤을 옆으로 눕혀 높이 559 mm (22 in) 바닥을 만들고 그 위에 비적재형 가전 19대를 세웁니다.

엔진에는 이 두 전략이 없었습니다. 둘 다 이제 결정론적 엔진 전략입니다. 이것이 뉴로-심볼릭 루프가 포착하도록 설계된 창의성입니다(5장).

4. 툴 암

두 번째 암에서는 모델이 `pack_cargo`라는 툴 하나로 엔진을 호출한 뒤 임의의 계획으로 답할 수 있다.

대부분의 모델은 매니페스트를 툴 호출에 복사해 넣고 엔진의 계획을 그대로 반환한다. 실패는 복사 오류뿐이다:

- Magistral이 잘못된 치수나 플래그를 가진 100개 항목을 보냈다;
- Kimi가 H10에서 잘못된 호출을 했다;
- Qwen이 4회 실행에서 실패했다.

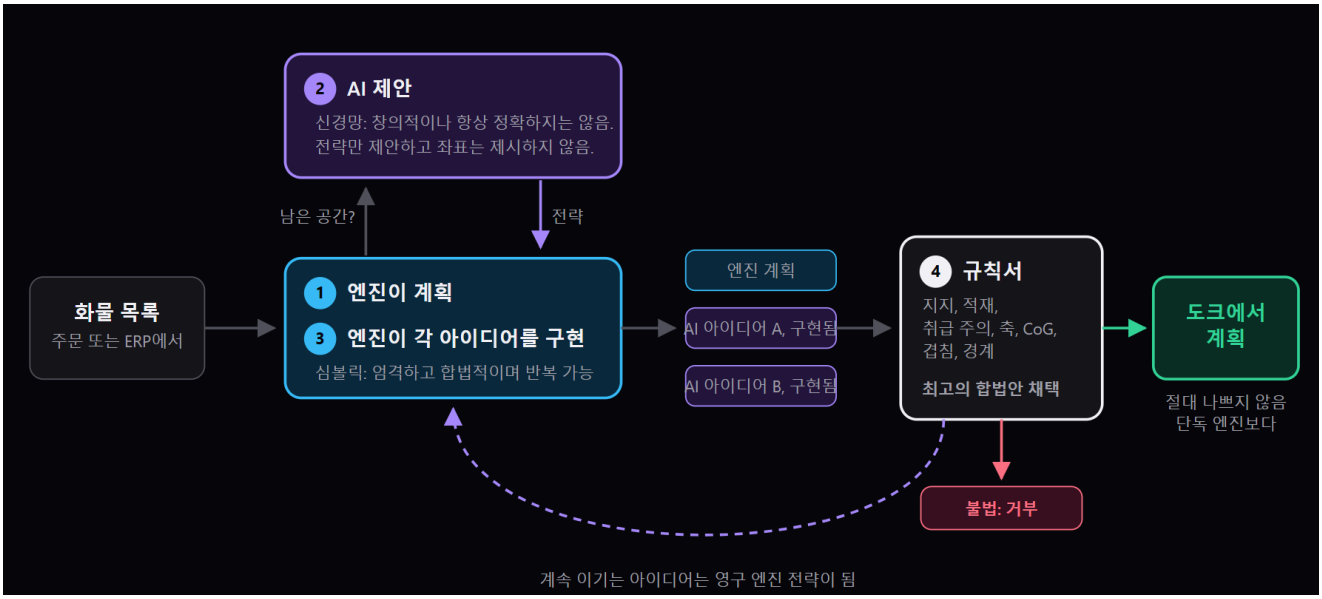
시사점: 어떤 결합 시스템에서도 엔진은 모델을 거치지 않고 소스 시스템에서 직접 매니페스트를 읽어야 한다.

5. 뉴로-심볼릭 아키텍처

5.1 두 가지 지능

	패킹 엔진 (심볼릭)	생성형 AI (신경망)
규칙	인코딩된 모든 규칙을 매번 따름	주어진 규칙을 어김
재현성	같은 적재, 같은 계획	같은 적재에도 다른 계획이 나올 수 있음
감사	모든 배치에 이유가 있음	감사하기 어려움
범위	가진 전략으로 제한됨	새로운 전략을 찾아냄
규모	253개 물품 처리	대형 매니페스트에서 많은 모델이 답을 내지 못함

5.2 루프



1. **기준선.** 엔진이 소스 시스템에서 직접 읽은 매니페스트로 계획을 세웁니다.
2. **트리거.** 기준선에 여지가 있을 때만 모델을 호출합니다: 배치되지 않은 물품, 낮은 적재율, 또는 미달한 클리어런스 목표.
3. **제안.** 모델은 좌표가 아닌 전략 파라미터를 반환합니다:
   - 구역;
   - 순서;
   - 방향 선호도;
   - 바닥 및 상단 구조.
4. **실행.** 엔진이 제안마다 하나의 계획을 만듭니다.
5. **판정.** 검증기가 최상의 적법한 계획을 선택합니다. 기준선은 항상 후보입니다.

6. **증류.** 반복해서 승리한 제안은 엔진의 영구 전략이 됩니다. 앞벽 및 카턴 바닥 전략이 그랬던 것처럼.

5.3 설계상 특성

- **엔진 단독보다 나쁘지 않음,** 기준선이 항상 후보이기 때문입니다.
- **부적법한 계획이 도크에 도달하지 않음,** 검증기가 모든 후보를 판정하기 때문입니다.
- **전사 위험 없음,** 모델이 매니페스트나 좌표를 건드리지 않기 때문입니다(4장).
- **모델 출력 제한으로 인한 실패 없음,** 제안이 253개 물품 계획이 아닌 몇 개의 파라미터이기 때문입니다(3.1장).
- **제한된 AI 비용,** 모델이 트리거에서만 실행되기 때문입니다.

우리는 지금 이 루프를 구축하고 있습니다. 다음 판에서는 계획당 비용과 시간으로 자체 참가자로 채점될 것입니다.

6. 범위와 한계

6.1 검증기 범위

적법은 v2.3.0에 정의된 규칙 기준의 적법을 의미합니다. 검증기는 다중 경유 시 후입선출 외의 적재 순서는 검사하지 않습니다.

6.2 편차

각 참가자는 케이스당 한 번씩 실행합니다.

- **엔진**은 시간 예산 내에서 전략들을 경쟁시키므로, 시스템 부하에 따라 먼저 완료되는 전략이 달라질 수 있습니다. H02에서 실행 간 5.63-5.79 m (18.5-19.0 ft)를 측정했습니다.
- **모델**은 한 번만 샘플링합니다.

단일 실행 간 2% 미만의 차이는 근소한 것으로 보고됩니다.

6.3 모델 티어 및 출력 제한

- 각 랩의 주류 프런티어 모델을 실행했습니다. 프리미엄 티어(Anthropic Fable, OpenAI GPT-6 Astra)는 실행하지 않았습니다.
- 출력 제한을 늘리면 일부 무응답 실행(섹션 3.1)이 계획으로 전환될 수 있습니다.

6.4 케이스 출처

모든 케이스는 한 플랫폼 사용자의 것입니다. 다른 화주의 화물 구성은 다를 수 있습니다.

7. 다음 단계

7.1 실무 규칙

다음으로 엔진과 validator v3에 실무 규칙 두 가지가 적용됩니다:

- **지붕 여유 공간:** 지붕 아래 최소 50 mm (2.0 in), 또는 25 kg (55 lb) 초과 적재 물품 위 100 mm (3.9 in).
- **측면 유격:** 포크리프트 열마다.

지붕 규칙으로 모든 참가자를 재채점했습니다. 기울일 수 없고 지붕에서 50 mm (2.0 in)를 뺀 높이보다 높은 바닥 적재 물품은 예외입니다. 해당 케이스가 강제하기 때문입니다.

- **엔진**은 표준 런 16개 중 9개, 대형 런 8개 중 1개에서 규칙을 지킵니다.
- **6개 런에서 모델 답변이 규칙을 지키며** 엔진에 가까운 계획을 세웁니다. 이들이 엔진의 첫 번째 목표입니다:
  - H02 (GPT);

    - H06 (Grok, 11.41 m / 37.4 ft);
    - H12 LDM (Gemini, 11.00 m / 36.1 ft);
    - H12 팔레트 (Kimi, 28 위치);
    - BV3 LDM (GPT, 9.14 m / 30.0 ft);
    - BV5 (Opus, 124 배치).
- **8개 런에서는 어떤 참가자도 규칙을 지키지 못했습니다:** H01, H03, H08, BV1, BV2 런 2개 모두, BV3 팔레트, BV6.

7.2 다음 에디션

- Validator v3: 지붕 여유 공간, 측면 유격, 적재 순서 검사 포함, 모든 참가자에게 적용.
- 새로운 실제 고객 케이스.
- 각 랩의 최고 성능 모델(프리미엄 티어 포함), 모델당 여러 차례 런 및 분산 보고.
- 뉴로-심볼릭 루프를 참가자로.

8. 공개

공개:

- 모든 케이스 전체;
- 검증기;
- 프롬프트;
- 채점 및 순위 코드;
- 모든 실행마다: 보낸 요청, 제공자의 원본 응답, 파싱된 계획.

뷰어: 모든 참가자의 모든 계획을 3D로, 해당 프롬프트 및 원본 답변과 함께: hansatic.com/ko/neuro-symbolic-load-planning.

비공개:

- 팔레트 위치 카운터(Hansatic의 운영 카운터, 모든 참가자에게 동일하게 적용);
- 엔진;
- 소스 계획의 ID.

부록 A. 프롬프트 템플릿

차량 유형당 시스템 프롬프트 하나. 모델에게 보낸 그대로입니다. 사용자 메시지는 먼저 목표를, 그다음 케이스를 JSON으로 줍니다. 항목 목록은 여기서 축약했습니다. 프롬프트는 영어로 표시됩니다. 모델이 받은 언어가 영어이기 때문입니다.

해상 컨테이너 (BV1, 40' 하이 큐브)

시스템 프롬프트

```
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"]}
```

사용자 메시지 (품목 목록 축약)

```
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"
 ]
}
```

트럭 (H09, 13.6 m (44.6 ft) 커튼사이더)

시스템 프롬프트

```
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"]}
```

사용자 메시지 (품목 목록 축약)

```
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"
```

```
 ]
}
```

항공 컨테이너 (U01, 경사진 벽이 있는 AKE)

시스템 프롬프트

```
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"]}
```

사용자 메시지 (품목 목록 축약)

```
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"
 ]
}
```


팔레트 (P01, 1200 × 1000 mm (47 × 39 in))

시스템 프롬프트

```
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"]}
```

사용자 메시지 (품목 목록 축약)

```
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"
```

```
  ]
}
```

부록 B. 전체 26개 런

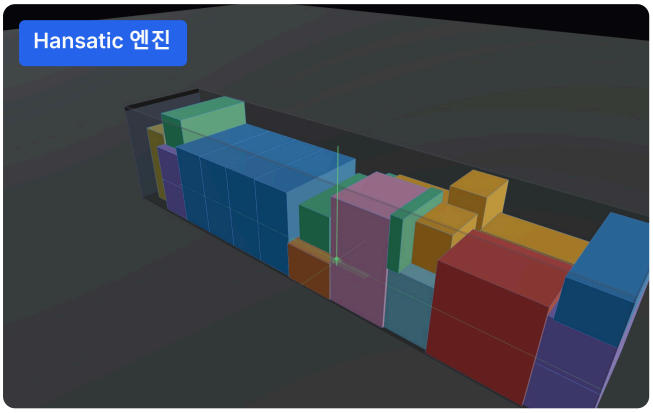
런별 한 페이지: 화물, 엔진의 계획과 최적의 적법한 모델 답변을 나란히, 그리고 모든 참가자의 결과.

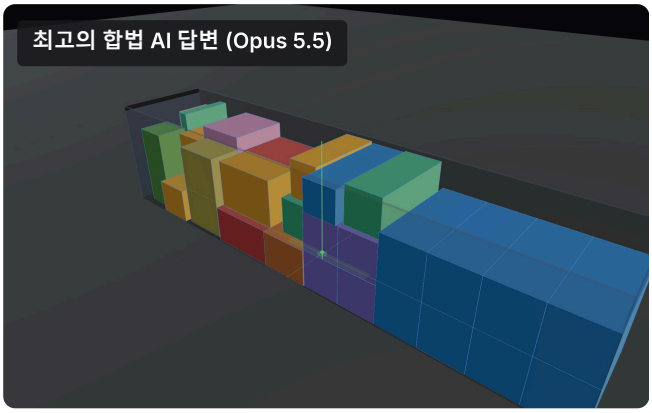
H01 최단 적재

차량	개수	목표	결과
40' High Cube	27	최단 적재	엔진 근소하게 앞섬

어려운 점

27개 세워 두어야 하는 품목





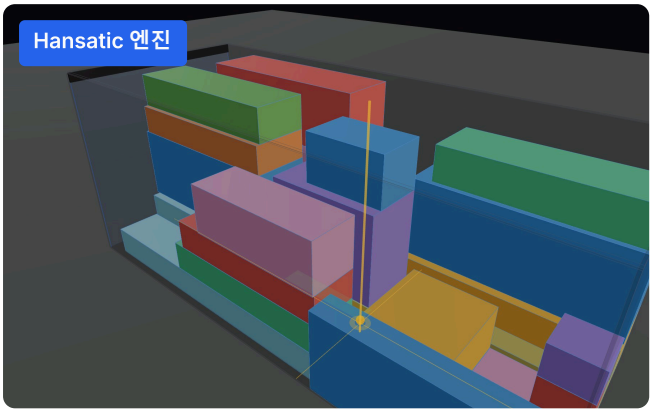
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**26 / 27**	**11.69 m (38.4 ft)**	
GPT-5.5	**합법**	25 / 27	11.88 m (39.0 ft)	
Opus 5.5	**합법**	26 / 27	11.74 m (38.5 ft)	
Grok 4.7	**합법**	25 / 27	11.37 m (37.3 ft)	
Kimi K3.1 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
DeepSeek V4 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Gemini 4 Argon	**불법**	22 / 27	11.56 m (37.9 ft)	1 잘못된 화물 치수, 1 차량 외부, 1 떠 있는 화물
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	**불법**	27 / 27	12.09 m (39.7 ft)	7 화물 겹침, 14 차량 외부, 3 높이 제한 초과, 1 떠 있는 화물, 1 무게 중심 이탈

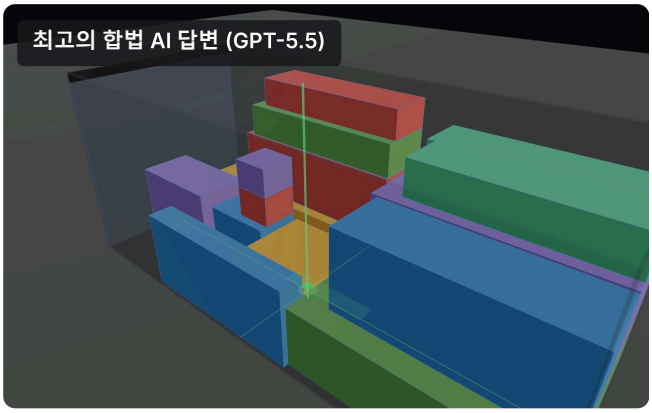
H02 최단 적재

차량	개수	목표	결과
20' Dry Container	29	최단 적재	엔진 확실히 앞섬

어려운 점

29개 세워 두어야 하는 품목; 최대 5.00 m (16.4 ft) 길이 품목; 가장 무거운 품목 1,550 kg (3,417 lb)





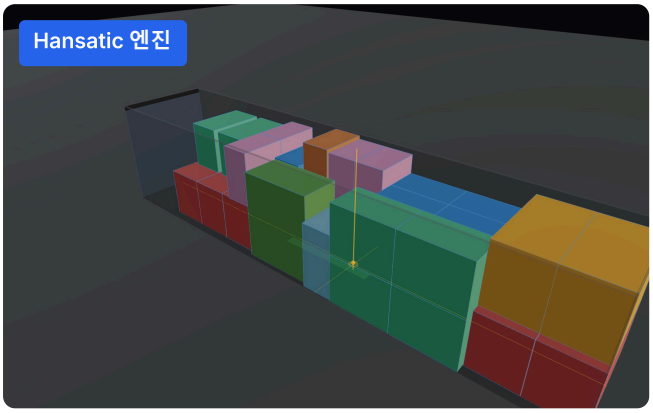
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**24 / 29**	**5.78 m (19.0 ft)**	
GPT-5.5	**합법**	18 / 29	5.76 m (18.9 ft)	
Opus 5.5	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Grok 4.7	제공자 오류	-	-	제공자가 요청을 거부함.
Kimi K3.1 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
DeepSeek V4 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Gemini 4 Argon	**불법**	23 / 29	5.00 m (16.4 ft)	6 잘못된 화물 치수, 13 화물 겹침, 9 떠 있는 화물, 9 찌그러진 화물
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	**불법**	29 / 29	7.97 m (26.1 ft)	37 화물 겹침, 23 차량 외부, 1 높이 제한 초과, 7 떠 있는 화물, 5 찌그러진 화물, 2 무게 중심 이탈

H03 최단 적재

차량	개수	목표	결과
40' High Cube	32	최단 적재	엔진 확실히 앞섬

어려운 점

25개 세워 두어야 하는 품목; 가장 무거운 품목 2,680 kg (5,908 lb)



이번 실행에서 합법적인 계획을 낸 AI 모델이 없습니다.

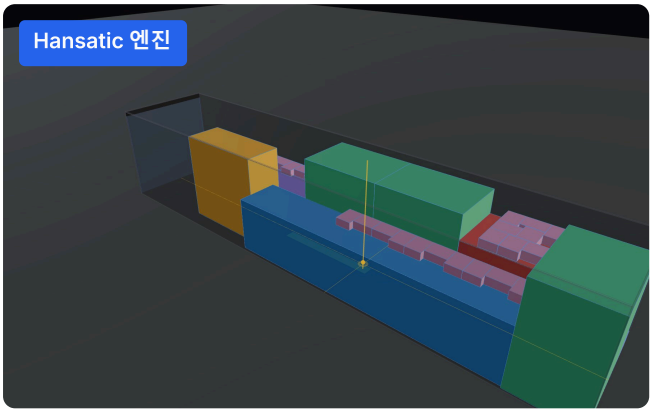
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**31 / 32**	**10.51 m (34.5 ft)**	
GPT-5.5	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Opus 5.5	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Grok 4.7	**불법**	31 / 32	12.03 m (39.5 ft)	1 찌그러진 화물
Kimi K3.1 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
DeepSeek V4 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Gemini 4 Argon	**불법**	31 / 32	11.81 m (38.7 ft)	1 잘못된 화물 치수, 12 화물 겹침, 6 차량 외부, 3 높이 제한 초과, 2 떠 있는 화물, 3 찌그러진 화물
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	계획 반환 없음	-	-	답변에 유효한 계획이 없었음.

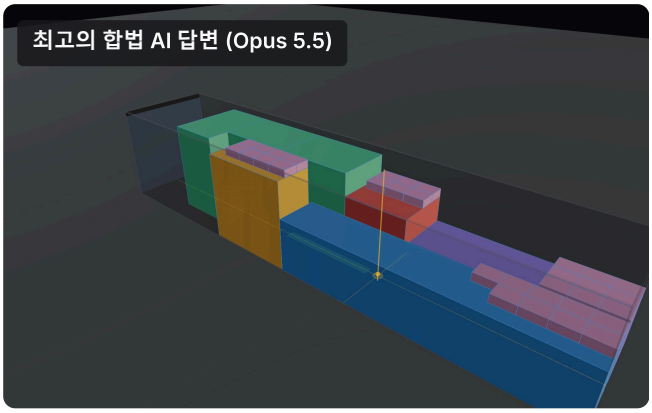
H04 최단 적재

차량	개수	목표	결과
40' Dry Container	39	최단 적재	동점

어려운 점

최대 6.30 m (20.7 ft) 길이 품목; 가장 무거운 품목 998 kg (2,200 lb)





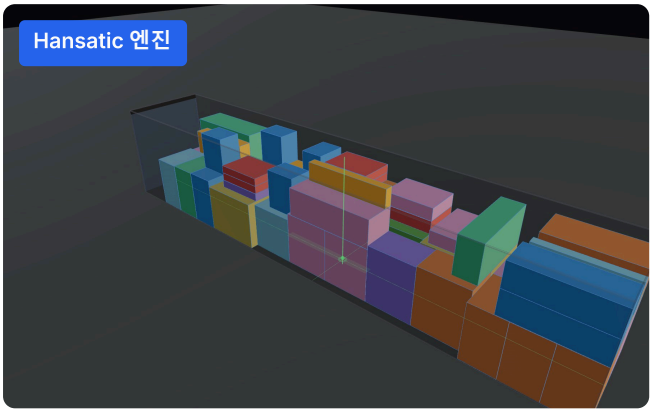
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**39 / 39**	**10.44 m (34.3 ft)**	
GPT-5.5	**합법**	39 / 39	10.72 m (35.2 ft)	
Opus 5.5	**합법**	39 / 39	10.44 m (34.3 ft)	
Grok 4.7	**합법**	39 / 39	11.10 m (36.4 ft)	
Kimi K3.1 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
DeepSeek V4 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Gemini 4 Argon	**불법**	29 / 39	6.30 m (20.7 ft)	10 잘못된 화물 치수, 6 떠 있는 화물
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	**불법**	38 / 39	10.20 m (33.5 ft)	1 잘못된 화물 치수, 27 화물 겹침, 6 차량 외부, 24 떠 있는 화물, 1 찌그러진 화물, 1 무게 중심 이탈

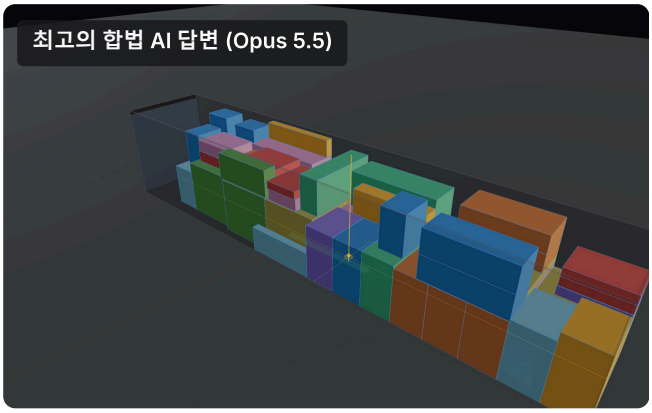
H05 최단 적재

차량	개수	목표	결과
45' High Cube	49	최단 적재	동점

어려운 점

49개 세워 두어야 하는 품목; 27개 적재 중량 제한 품목; 가장 무거운 품목 814 kg (1,795 lb)





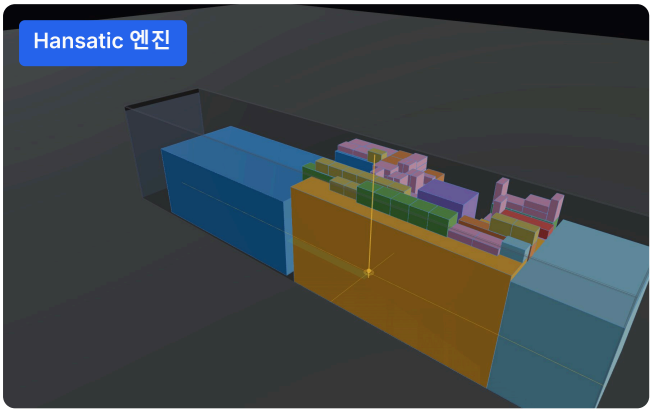
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**49 / 49**	**12.40 m (40.7 ft)**	
GPT-5.5	**합법**	45 / 49	13.28 m (43.6 ft)	
Opus 5.5	**합법**	49 / 49	12.40 m (40.7 ft)	
Grok 4.7	**합법**	43 / 49	12.96 m (42.5 ft)	
Kimi K3.1 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
DeepSeek V4 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Gemini 4 Argon	**불법**	37 / 49	10.49 m (34.4 ft)	11 잘못된 화물 치수, 2 화물 겹침, 5 차량 외부, 2 높이 제한 초과, 3 떠 있는 화물, 2 찌그러진 화물
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	**불법**	47 / 49	12.73 m (41.8 ft)	2 잘못된 화물 치수, 31 화물 겹침, 9 차량 외부, 14 높이 제한 초과, 21 떠 있는 화물, 2 찌그러진 화물

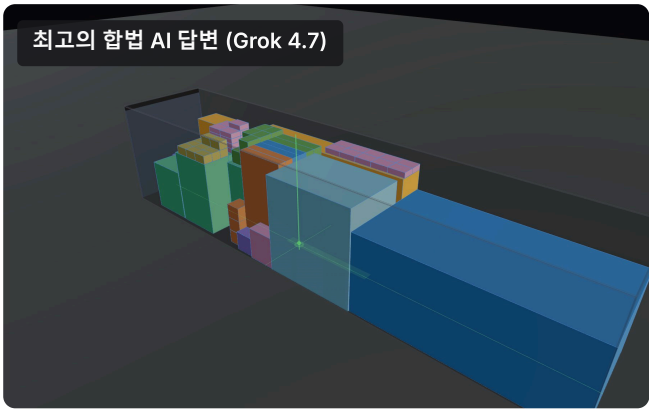
H06 최단 적재

차량	개수	목표	결과
40' High Cube	83	최단 적재	엔진 확실히 앞섬

어려운 점

5개 적재 불가 품목; 1개 세워 두어야 하는 품목; 가장 무거운 품목 3,020 kg (6,658 lb)





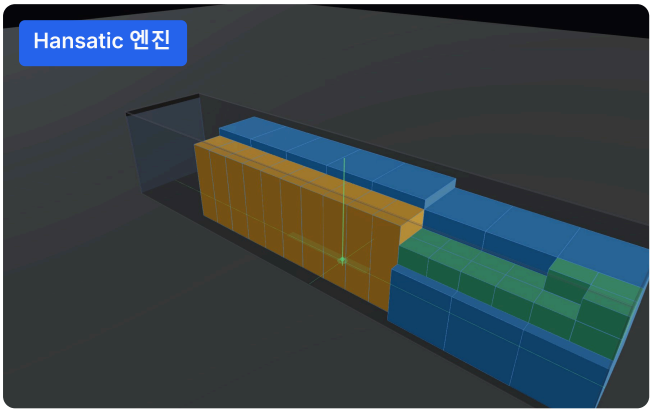
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**83 / 83**	**10.65 m (35.0 ft)**	
GPT-5.5	**합법**	83 / 83	11.78 m (38.7 ft)	
Opus 5.5	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Grok 4.7	**합법**	83 / 83	11.41 m (37.4 ft)	
Kimi K3.1 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
DeepSeek V4 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Gemini 4 Argon	**불법**	7 / 83	6.89 m (22.6 ft)	2 잘못된 화물 치수, 1 무게 중심 이탈
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	**불법**	20 / 83	12.25 m (40.2 ft)	2 화물 겹침, 4 차량 외부, 1 높이 제한 초과, 14 떠 있는 화물, 1 무게 중심 이탈

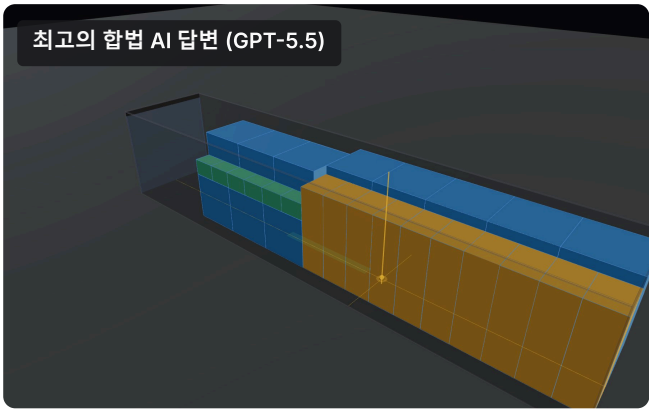
H07 최단 적재

차량	개수	목표	결과
40' Dry Container	45	최단 적재	동점

어려운 점

10개 적재 불가 품목; 15개 세워 두어야 하는 품목; 15개 적재 중량 제한 품목; 가장 무거운 품목 1,200 kg (2,646 lb)





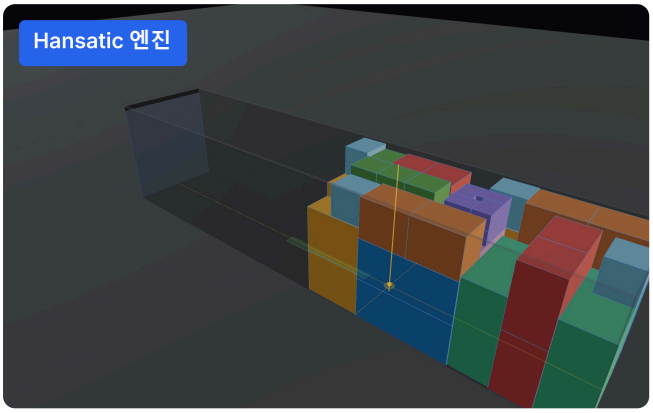
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**45 / 45**	**9.60 m (31.5 ft)**	
GPT-5.5	합법	45 / 45	9.60 m (31.5 ft)	
Opus 5.5	합법	45 / 45	9.61 m (31.5 ft)	
Grok 4.7	불법	45 / 45	10.80 m (35.4 ft)	1 무게 중심 이탈
Kimi K3.1 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
DeepSeek V4 Pro	합법	45 / 45	12.00 m (39.4 ft)	
Gemini 4 Argon	불법	32 / 45	7.60 m (24.9 ft)	13 잘못된 화물 치수, 1 화물 겹침, 16 떠 있는 화물
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	불법	10 / 45	12.00 m (39.4 ft)	1 무게 중심 이탈

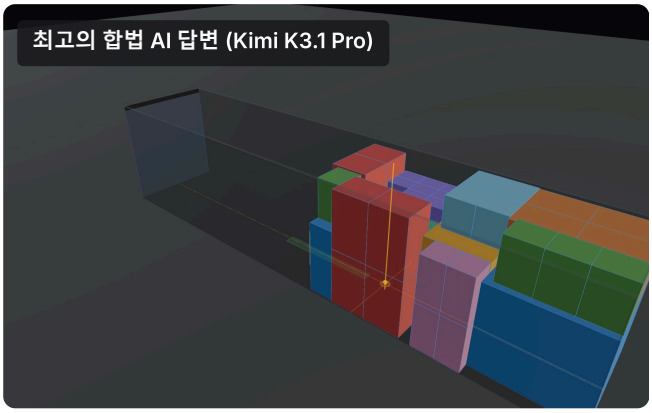
H08 최단 적재

차량	개수	목표	결과
40' High Cube	52	최단 적재	엔진 확실히 앞섬

어려운 점

52개 세워 두어야 하는 품목





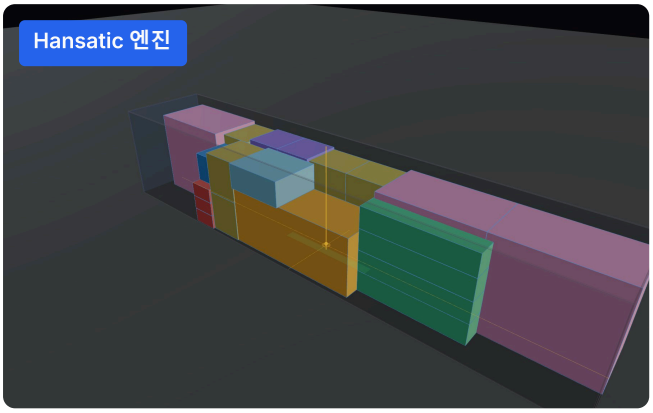
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**52 / 52**	**5.92 m (19.4 ft)**	
GPT-5.5	**합법**	52 / 52	6.26 m (20.5 ft)	
Opus 5.5	**합법**	52 / 52	6.20 m (20.3 ft)	
Grok 4.7	**합법**	52 / 52	6.18 m (20.3 ft)	
Kimi K3.1 Pro	**합법**	52 / 52	6.12 m (20.1 ft)	
DeepSeek V4 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Gemini 4 Argon	**불법**	51 / 52	6.51 m (21.4 ft)	1 잘못된 화물 치수, 4 화물 겹침, 1 떠 있는 화물
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	**불법**	52 / 52	11.25 m (36.9 ft)	14 화물 겹침, 20 차량 외부

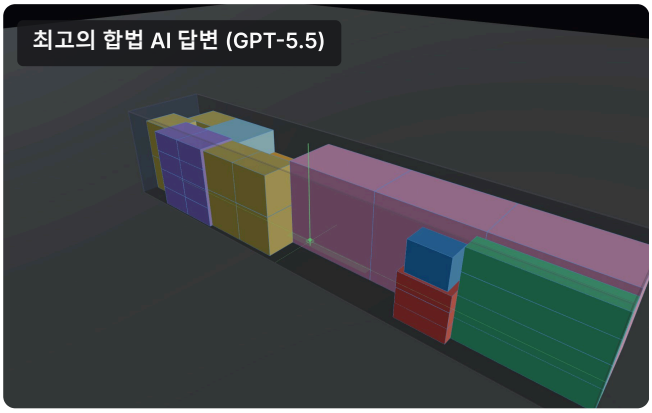
H09 최단 적재

차량	개수	목표	결과
13.6m Curtainsider	30	최단 적재	엔진 근소하게 앞섬

어려운 점

30개 세워 두어야 하는 품목; 가장 무거운 품목 2,400 kg (5,291 lb)





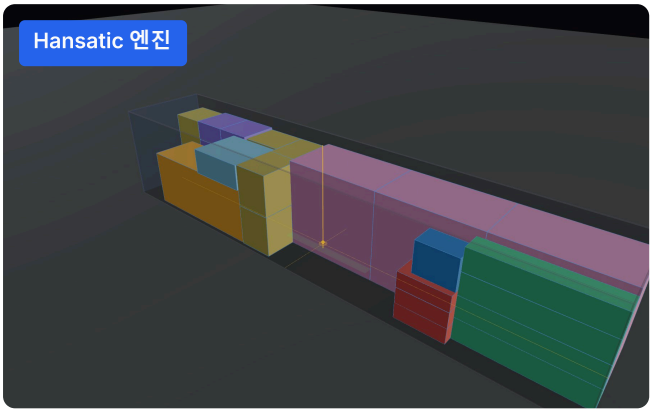
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**29 / 30**	**13.36 m (43.8 ft)**	
GPT-5.5	**합법**	29 / 30	13.56 m (44.5 ft)	
Opus 5.5	**합법**	29 / 30	13.60 m (44.6 ft)	
Grok 4.7	**합법**	29 / 30	13.60 m (44.6 ft)	
Kimi K3.1 Pro	**합법**	29 / 30	13.60 m (44.6 ft)	
DeepSeek V4 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Gemini 4 Argon	**불법**	19 / 30	11.60 m (38.1 ft)	11 잘못된 화물 치수, 7 화물 겹침
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	**불법**	30 / 30	13.36 m (43.8 ft)	17 화물 겹침, 25 차량 외부, 7 떠 있는 화물, 1 축 하중 초과

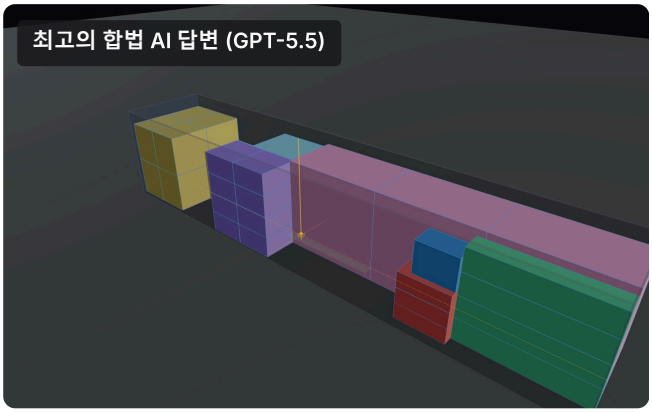
H09 최소 팔레트 위치

차량	개수	목표	결과
13.6m Curtainsider	30	최소 팔레트 위치	엔진 근소하게 앞섬

어려운 점

30개 세워 두어야 하는 품목; 가장 무거운 품목 2,400 kg (5,291 lb)





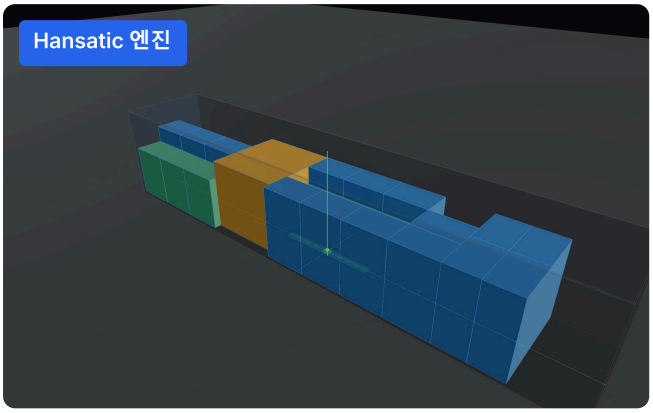
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**29 / 30**	**29개 위치, 13.36 m (43.8 ft)**	
GPT-5.5	**합법**	29 / 30	29개 위치, 13.60 m (44.6 ft)	
Opus 5.5	**합법**	29 / 30	31개 위치, 13.60 m (44.6 ft)	
Grok 4.7	**합법**	29 / 30	29개 위치, 13.62 m (44.7 ft)	
Kimi K3.1 Pro	**합법**	29 / 30	29개 위치, 13.60 m (44.6 ft)	
DeepSeek V4 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Gemini 4 Argon	**불법**	30 / 30	29개 위치, 11.72 m (38.5 ft)	18 화물 겹침, 1 차량 외부, 3 찌그러진 화물, 1 축 하중 초과
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	**불법**	30 / 30	34개 위치, 13.36 m (43.8 ft)	4 화물 겹침, 25 차량 외부, 1 축 하중 초과, 1 정렬 어긋난 열

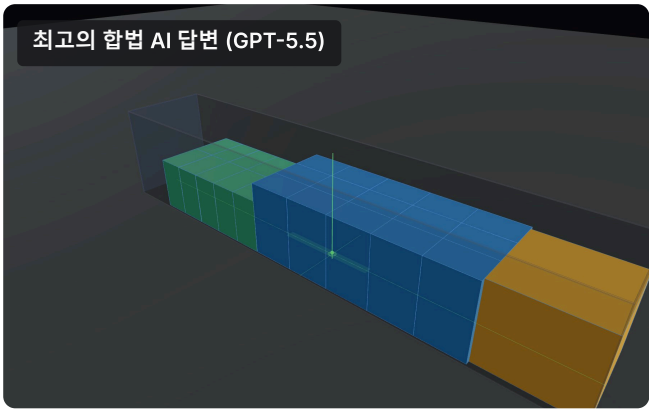
H10 최단 적재

차량	개수	목표	결과
13.6m Curtainsider	26	최단 적재	동점

어려운 점

26개 적재 불가 품목; 16개 세워 두어야 하는 품목; 가장 무거운 품목 666 kg (1,468 lb)





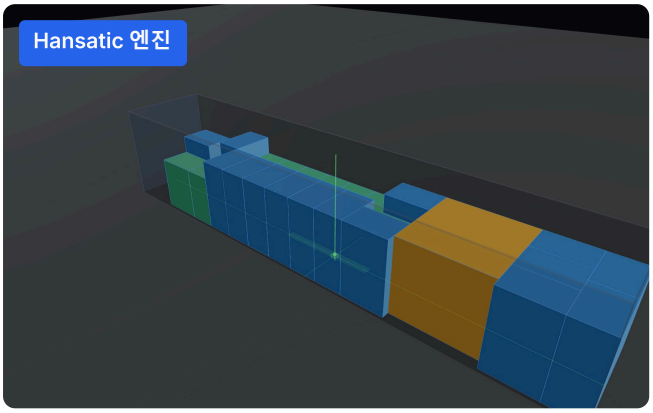
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**26 / 26**	**12.20 m (40.0 ft)**	
GPT-5.5	합법	26 / 26	12.20 m (40.0 ft)	
Opus 5.5	합법	26 / 26	12.20 m (40.0 ft)	
Grok 4.7	합법	26 / 26	12.20 m (40.0 ft)	
Kimi K3.1 Pro	합법	26 / 26	12.20 m (40.0 ft)	
DeepSeek V4 Pro	합법	26 / 26	12.20 m (40.0 ft)	
Gemini 4 Argon	합법	26 / 26	12.60 m (41.3 ft)	
Qwen3p8	합법	26 / 26	12.20 m (40.0 ft)	
Magistral Medium	합법	26 / 26	13.00 m (42.7 ft)	

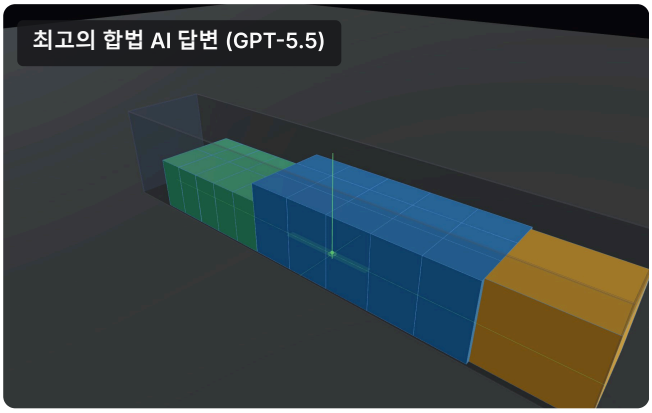
H10 최소 팔레트 위치

차량	개수	목표	결과
13.6m Curtainsider	26	최소 팔레트 위치	동점

어려운 점

26개 적재 불가 품목; 16개 세워 두어야 하는 품목; 가장 무거운 품목 666 kg (1,468 lb)





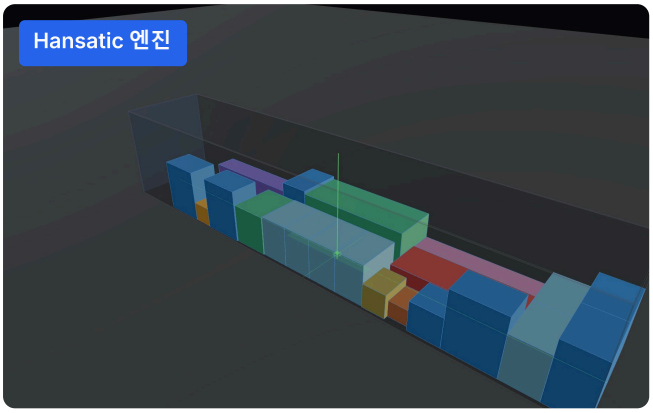
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**26 / 26**	**31개 위치, 12.20 m (40.0 ft)**	
GPT-5.5	합법	26 / 26	31개 위치, 12.20 m (40.0 ft)	
Opus 5.5	합법	26 / 26	31개 위치, 13.00 m (42.7 ft)	
Grok 4.7	합법	26 / 26	31개 위치, 12.20 m (40.0 ft)	
Kimi K3.1 Pro	합법	26 / 26	31개 위치, 12.20 m (40.0 ft)	
DeepSeek V4 Pro	합법	26 / 26	31개 위치, 12.20 m (40.0 ft)	
Gemini 4 Argon	불법	26 / 26	31개 위치, 12.60 m (41.3 ft)	1 축 하중 초과
Qwen3p8	합법	26 / 26	31개 위치, 13.00 m (42.7 ft)	
Magistral Medium	불법	26 / 26	28개 위치, 13.00 m (42.7 ft)	10 차량 외부

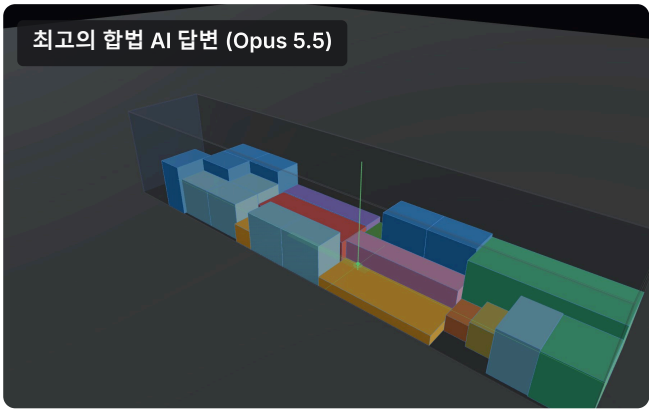
H11 최단 적재

차량	개수	목표	결과
13.6m Curtainsider	23	최단 적재	엔진 근소하게 앞섬

어려운 점

23개 적재 불가 품목; 23개 세워 두어야 하는 품목; 가장 무거운 품목 2,700 kg (5,952 lb)





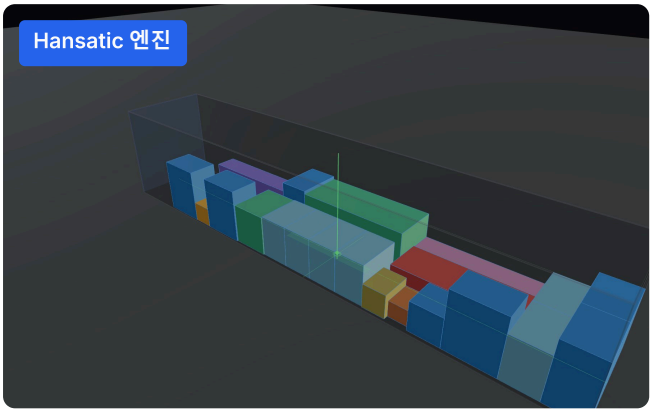
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**23 / 23**	**12.00 m (39.4 ft)**	
GPT-5.5	**합법**	23 / 23	12.42 m (40.7 ft)	
Opus 5.5	**합법**	23 / 23	12.20 m (40.0 ft)	
Grok 4.7	**합법**	23 / 23	12.62 m (41.4 ft)	
Kimi K3.1 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
DeepSeek V4 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Gemini 4 Argon	**합법**	23 / 23	13.02 m (42.7 ft)	
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	**불법**	23 / 23	12.34 m (40.5 ft)	7 화물 겹침, 8 차량 외부

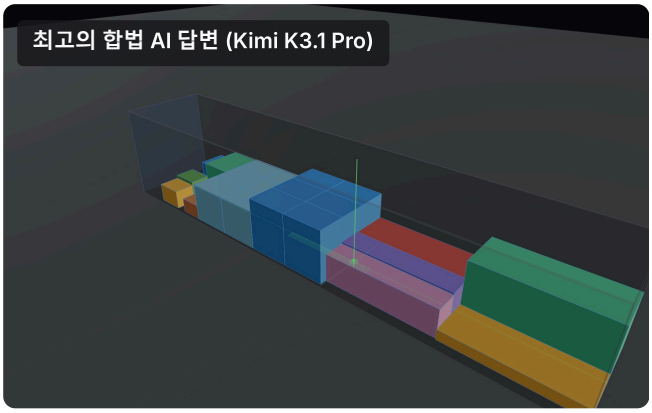
H11 최소 팔레트 위치

차량	개수	목표	결과
13.6m Curtainsider	23	최소 팔레트 위치	엔진 확실히 앞섬

어려운 점

23개 적재 불가 품목; 23개 세워 두어야 하는 품목; 가장 무거운 품목 2,700 kg (5,952 lb)





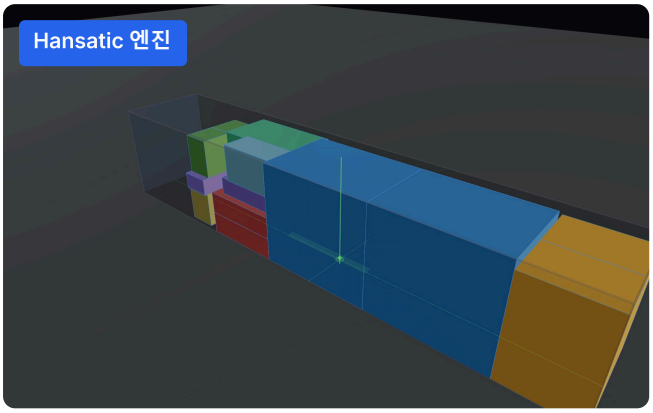
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**23 / 23**	**30개 위치, 12.00 m (39.4 ft)**	
GPT-5.5	**합법**	23 / 23	32개 위치, 12.64 m (41.5 ft)	
Opus 5.5	**합법**	23 / 23	32개 위치, 12.84 m (42.1 ft)	
Grok 4.7	제공자 오류	-	-	제공자가 요청을 거부함.
Kimi K3.1 Pro	**합법**	23 / 23	32개 위치, 12.62 m (41.4 ft)	
DeepSeek V4 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Gemini 4 Argon	**불법**	22 / 23	30개 위치, 12.44 m (40.8 ft)	1 잘못된 화물 치수
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	**불법**	23 / 23	30개 위치, 12.60 m (41.3 ft)	11 화물 겹침, 9 차량 외부

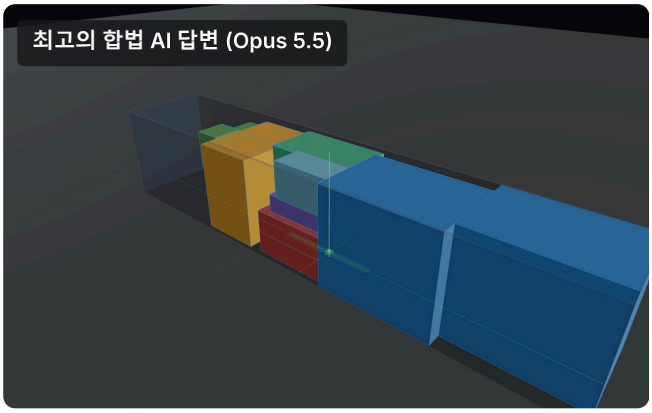
H12 최단 적재

차량	개수	목표	결과
13.6m Curtainsider	14	최단 적재	동점

어려운 점

6개 적재 불가 품목; 9개 세워 두어야 하는 품목; 가장 무거운 품목 732 kg (1,614 lb)





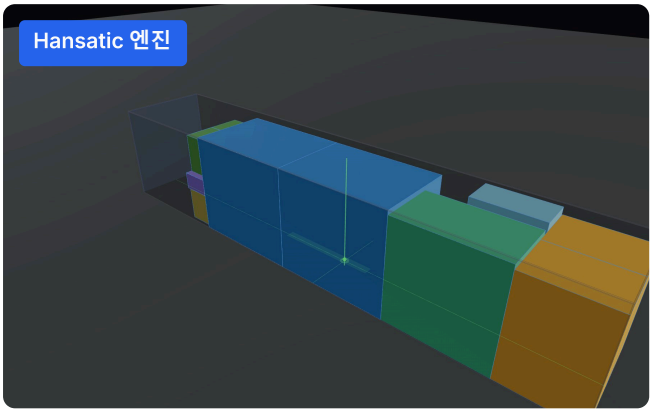
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**13 / 14**	**11.00 m (36.1 ft)**	
GPT-5.5	**합법**	13 / 14	11.40 m (37.4 ft)	
Opus 5.5	**합법**	13 / 14	11.00 m (36.1 ft)	
Grok 4.7	제공자 오류	-	-	제공자가 요청을 거부함.
Kimi K3.1 Pro	**합법**	13 / 14	11.00 m (36.1 ft)	
DeepSeek V4 Pro	**합법**	13 / 14	11.70 m (38.4 ft)	
Gemini 4 Argon	**합법**	13 / 14	11.00 m (36.1 ft)	
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	**불법**	14 / 14	13.50 m (44.3 ft)	5 화물 겹침, 5 차량 외부, 2 떠 있는 화물

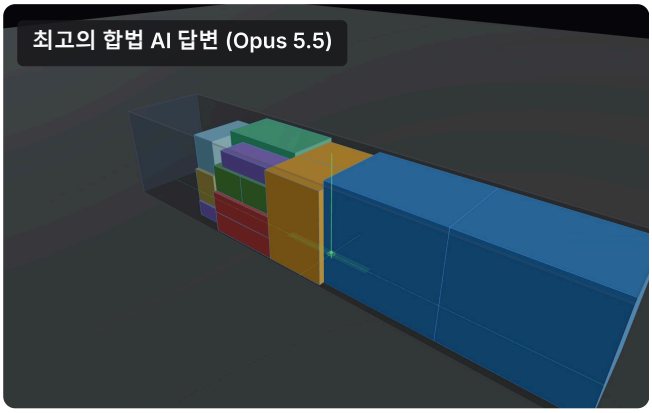
H12 최소 팔레트 위치

차량	개수	목표	결과
13.6m Curtainsider	14	최소 팔레트 위치	엔진 확실히 앞섬

어려운 점

6개 적재 불가 품목; 9개 세워 두어야 하는 품목; 가장 무거운 품목 732 kg (1,614 lb)





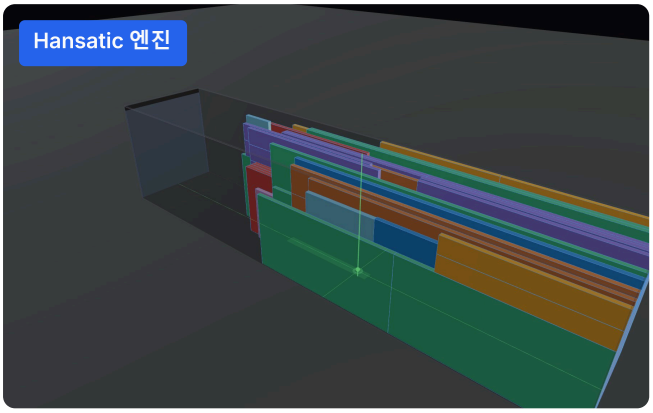
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**13 / 14**	**25개 위치, 11.00 m (36.1 ft)**	
GPT-5.5	**합법**	13 / 14	29개 위치, 11.73 m (38.5 ft)	
Opus 5.5	**합법**	13 / 14	28개 위치, 11.00 m (36.1 ft)	
Grok 4.7	제공자 오류	-	-	제공자가 요청을 거부함.
Kimi K3.1 Pro	**합법**	13 / 14	28개 위치, 11.00 m (36.1 ft)	
DeepSeek V4 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Gemini 4 Argon	**합법**	13 / 14	28개 위치, 11.15 m (36.6 ft)	
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	**불법**	14 / 14	33개 위치, 13.20 m (43.3 ft)	9 화물 겹침, 9 차량 외부

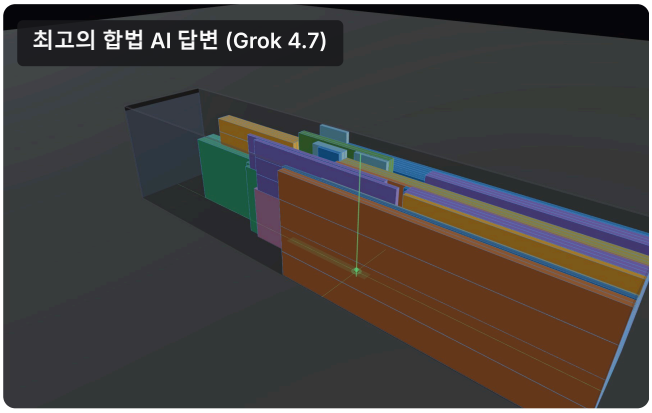
BV1 최단 적재

차량	개수	목표	결과
40' High Cube	103	최단 적재	엔진 확실히 앞섬

어려운 점

37개 세워 두어야 하는 품목; 최대 8.27 m (27.1 ft) 길이 품목





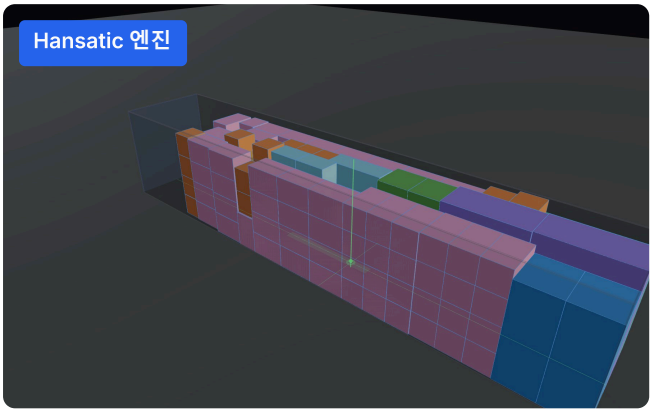
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**103 / 103**	**9.17 m (30.1 ft)**	
GPT-5.5	**합법**	92 / 103	8.27 m (27.1 ft)	
Opus 5.5	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Grok 4.7	**합법**	103 / 103	10.66 m (35.0 ft)	
Kimi K3.1 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
DeepSeek V4 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Gemini 4 Argon	**불법**	77 / 103	8.36 m (27.4 ft)	1 잘못된 화물 치수, 4 높이 제한 초과, 4 떠 있는 화물
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	**불법**	22 / 103	16.54 m (54.3 ft)	9 차량 외부

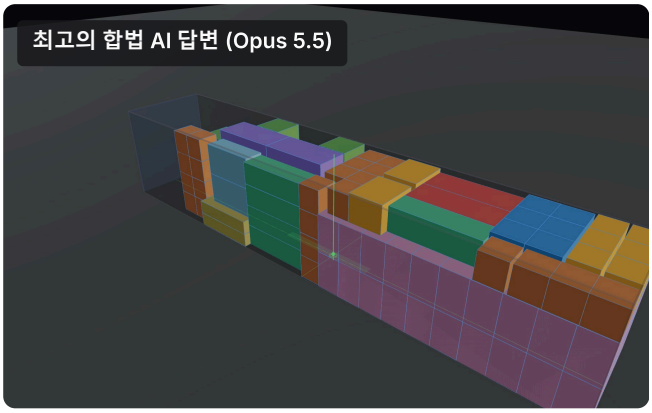
BV2 최단 적재

차량	개수	목표	결과
13.6m Curtainsider	153	최단 적재	엔진 확실히 앞섬

어려운 점

2개 적재 불가 품목; 가장 무거운 품목 578 kg (1,274 lb)





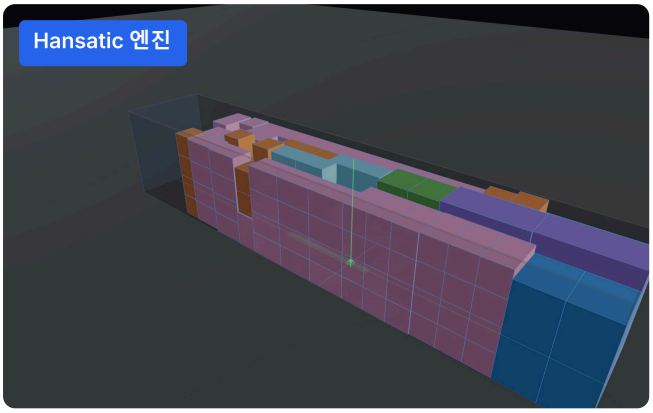
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**153 / 153**	**11.52 m (37.8 ft)**	
GPT-5.5	합법	153 / 153	13.14 m (43.1 ft)	
Opus 5.5	합법	153 / 153	11.78 m (38.6 ft)	
Grok 4.7	합법	153 / 153	12.30 m (40.4 ft)	
Kimi K3.1 Pro	합법	153 / 153	11.99 m (39.3 ft)	
DeepSeek V4 Pro	합법	153 / 153	12.79 m (42.0 ft)	
Gemini 4 Argon	불법	147 / 153	12.77 m (41.9 ft)	6 잘못된 화물 치수
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	불법	56 / 153	14.02 m (46.0 ft)	39 화물 겹침, 27 차량 외부

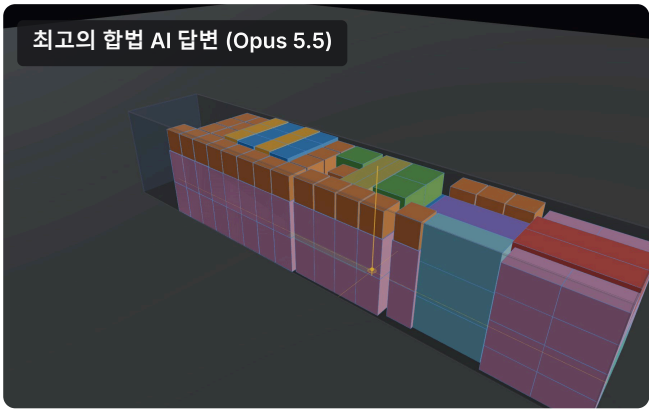
BV2 최소 팔레트 위치

차량	개수	목표	결과
13.6m Curtainsider	153	최소 팔레트 위치	엔진 근소하게 앞섬

어려운 점

2개 적재 불가 품목; 가장 무거운 품목 578 kg (1,274 lb)





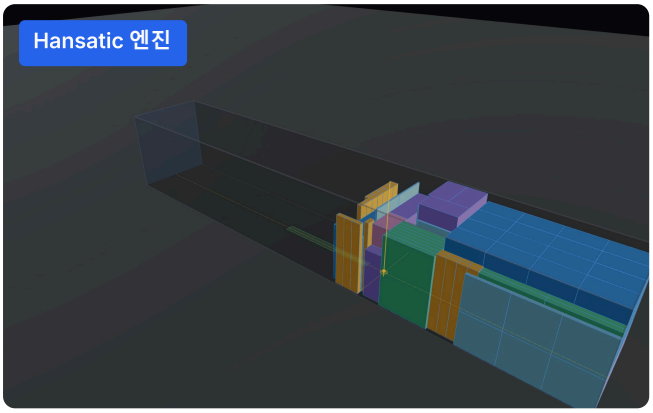
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**153 / 153**	**29개 위치, 11.52 m (37.8 ft)**	
GPT-5.5	**합법**	153 / 153	32개 위치, 12.60 m (41.3 ft)	
Opus 5.5	**합법**	153 / 153	29개 위치, 11.67 m (38.3 ft)	
Grok 4.7	**합법**	153 / 153	32개 위치, 12.78 m (41.9 ft)	
Kimi K3.1 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
DeepSeek V4 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Gemini 4 Argon	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	**불법**	153 / 153	34개 위치, 14.82 m (48.6 ft)	7 화물 겹침, 31 차량 외부, 86 높이 제한 초과, 96 떠 있는 화물

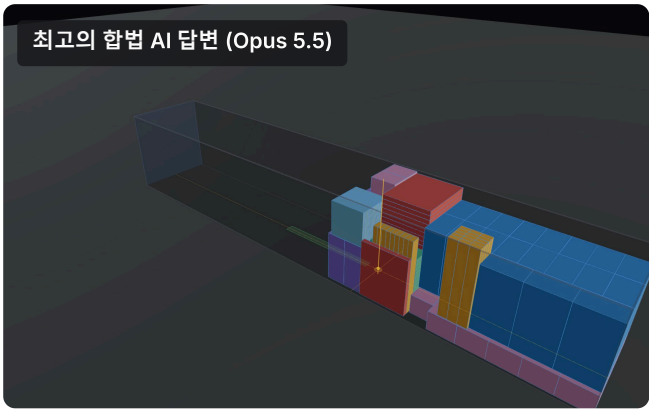
BV3 최단 적재

차량	개수	목표	결과
53ft Dry Van	119	최단 적재	엔진 확실히 앞섬

어려운 점

19개 적재 불가 품목; 35개 세워 두어야 하는 품목





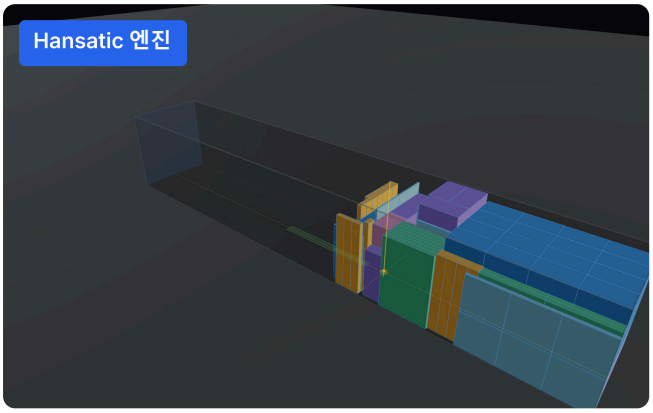
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**119 / 119**	**6.83 m (22.4 ft)**	
GPT-5.5	**합법**	119 / 119	9.14 m (30.0 ft)	
Opus 5.5	**합법**	119 / 119	6.99 m (22.9 ft)	
Grok 4.7	**합법**	119 / 119	8.10 m (26.6 ft)	
Kimi K3.1 Pro	계획 반환 없음	-	-	답변에 유효한 계획이 없었음.
DeepSeek V4 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Gemini 4 Argon	**불법**	115 / 119	8.36 m (27.4 ft)	4 잘못된 화물 치수, 5 화물 겹침, 3 찌그러진 화물
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	**불법**	22 / 119	16.76 m (55.0 ft)	11 차량 외부

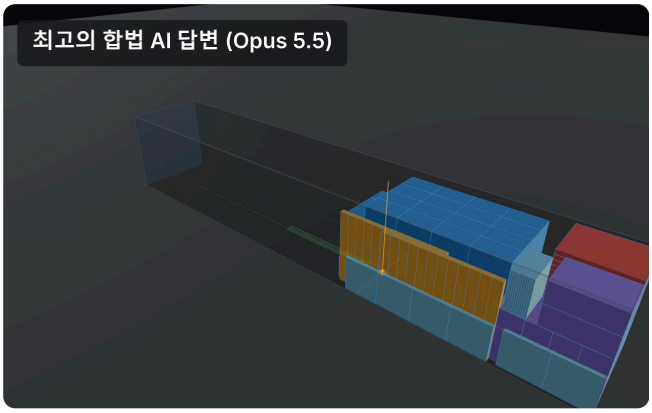
BV3 최소 팔레트 위치

차량	개수	목표	결과
53ft Dry Van	119	최소 팔레트 위치	엔진 확실히 앞섬

어려운 점

19개 적재 불가 품목; 35개 세워 두어야 하는 품목





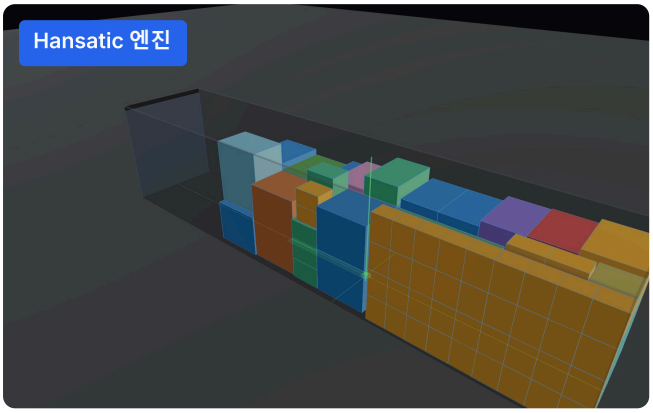
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**119 / 119**	**17개 위치, 6.83 m (22.4 ft)**	
GPT-5.5	**합법**	119 / 119	23개 위치, 9.07 m (29.8 ft)	
Opus 5.5	**합법**	119 / 119	18개 위치, 6.99 m (22.9 ft)	
Grok 4.7	**합법**	119 / 119	22개 위치, 8.64 m (28.3 ft)	
Kimi K3.1 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
DeepSeek V4 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Gemini 4 Argon	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	**불법**	107 / 119	21개 위치, 16.18 m (53.1 ft)	12 잘못된 화물 치수, 66 화물 겹침, 35 차량 외부, 60 떠 있는 화물, 1 찌그러진 화물, 1 고정되지 않은 고립 군집

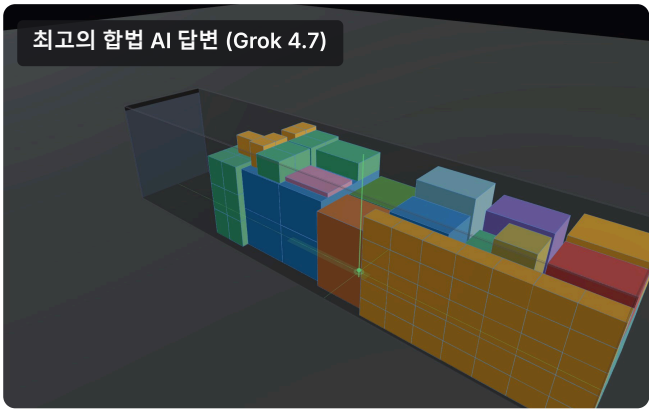
BV4 최단 적재

차량	개수	목표	결과
40' High Cube	145	최단 적재	엔진 확실히 앞섬

어려운 점

11개 적재 불가 품목; 11개 취급 주의 품목, 상단에만; 20개 세워 두어야 하는 품목





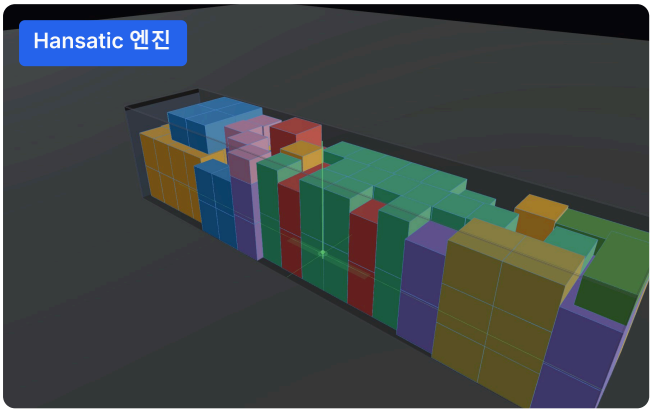
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**145 / 145**	**8.57 m (28.1 ft)**	
GPT-5.5	**합법**	145 / 145	11.96 m (39.2 ft)	
Opus 5.5	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Grok 4.7	**합법**	145 / 145	8.85 m (29.0 ft)	
Kimi K3.1 Pro	**합법**	145 / 145	10.87 m (35.7 ft)	
DeepSeek V4 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Gemini 4 Argon	**불법**	145 / 145	10.05 m (33.0 ft)	19 화물 겹침, 3 차량 외부, 9 높이 제한 초과, 4 떠 있는 화물
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	**합법**	11 / 145	10.33 m (33.9 ft)	

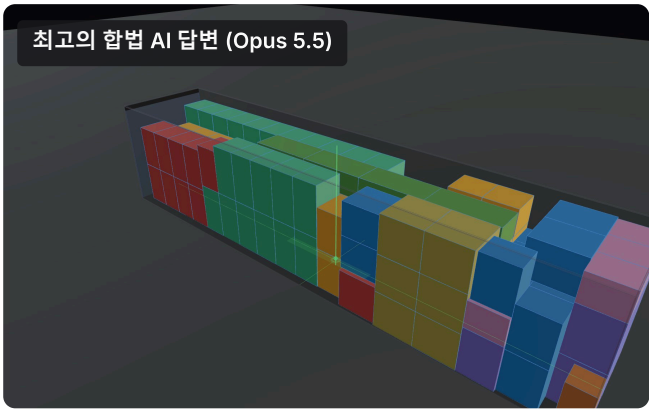
BV5 최단 적재

차량	개수	목표	결과
40' High Cube	136	최단 적재	엔진 확실히 앞섬

어려운 점

136개 세워 두어야 하는 품목





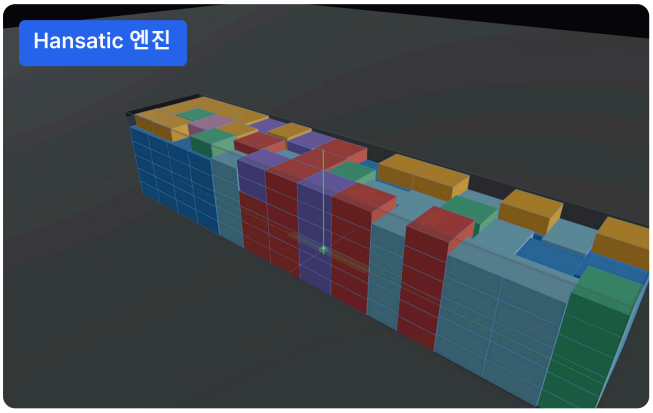
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**127 / 136**	**12.00 m (39.4 ft)**	
GPT-5.5	**합법**	115 / 136	11.73 m (38.5 ft)	
Opus 5.5	**합법**	124 / 136	11.77 m (38.6 ft)	
Grok 4.7	**합법**	115 / 136	11.87 m (39.0 ft)	
Kimi K3.1 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
DeepSeek V4 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Gemini 4 Argon	**불법**	121 / 136	12.00 m (39.4 ft)	1 기립 필수 화물 기울임, 12 화물 겹침, 6 차량 외부, 6 떠 있는 화물
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	**불법**	51 / 136	12.89 m (42.3 ft)	34 차량 외부, 6 떠 있는 화물

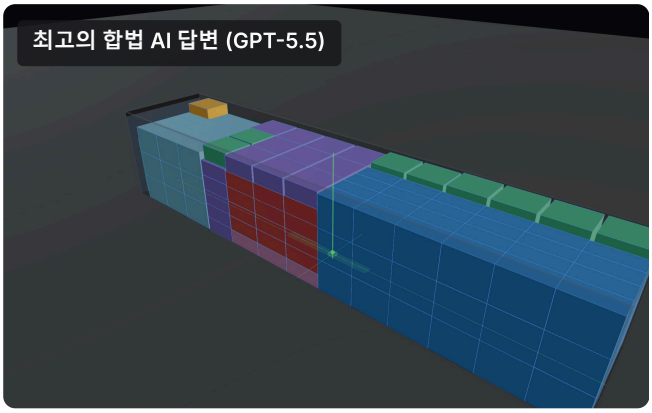
BV6 최단 적재

차량	개수	목표	결과
40' Dry Container	253	최단 적재	엔진 확실히 앞섬

어려운 점

다양한 크기, 특수 표시 없음



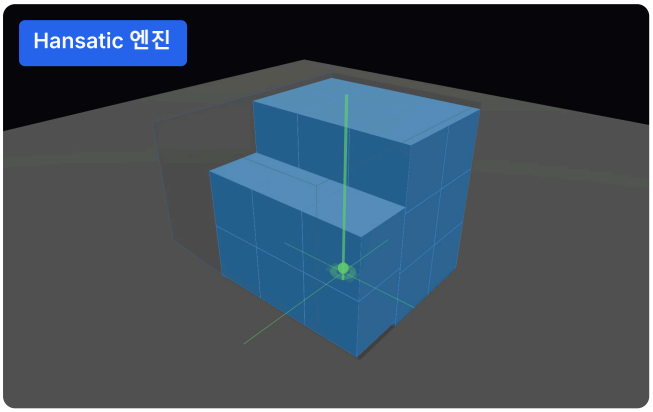


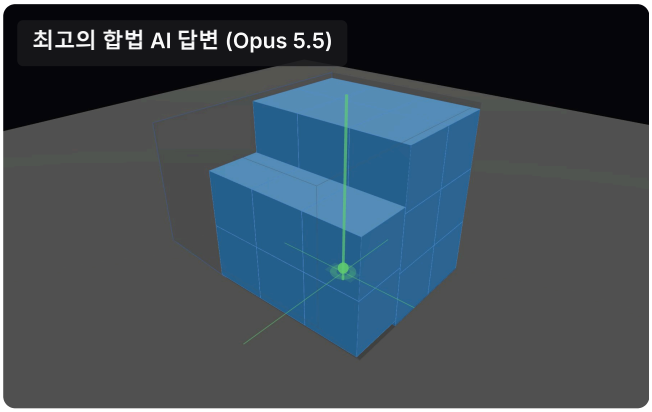
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**244 / 253**	**11.98 m (39.3 ft)**	
GPT-5.5	**합법**	238 / 253	11.97 m (39.3 ft)	
Opus 5.5	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Grok 4.7	**합법**	218 / 253	11.98 m (39.3 ft)	
Kimi K3.1 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
DeepSeek V4 Pro	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Gemini 4 Argon	**불법**	197 / 253	11.50 m (37.7 ft)	13 화물 겹침
Qwen3p8	계획 반환 없음	-	-	모델이 계획 작성 전에 출력 한도에 도달함.
Magistral Medium	**불법**	42 / 253	12.18 m (40.0 ft)	3 차량 외부

U01 최다 적재 개

차량	개수	목표	결과
AKE (LD3)	50	최다 적재 개	동점

어려운 점
경사진 벽





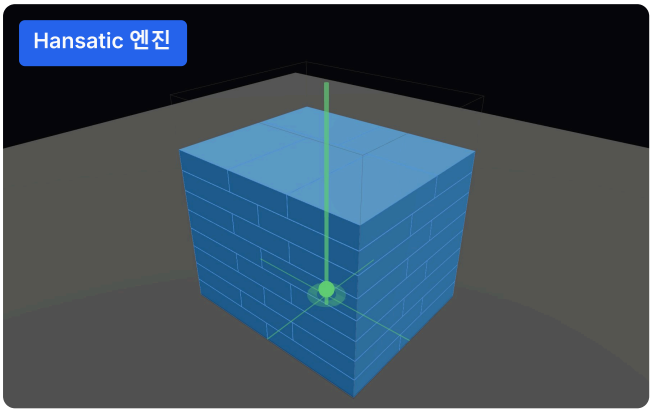
참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**24 / 50**	**24**	
GPT-5.5	합법	18 / 50	18	
Opus 5.5	합법	24 / 50	24	
Grok 4.7	합법	24 / 50	24	
Kimi K3.1 Pro	합법	24 / 50	24	
DeepSeek V4 Pro	합법	18 / 50	18	
Gemini 4 Argon	합법	24 / 50	24	
Qwen3p8	합법	24 / 50	24	
Magistral Medium	불법	24 / 50	24	3 항공 컨테이너 윤곽 밖

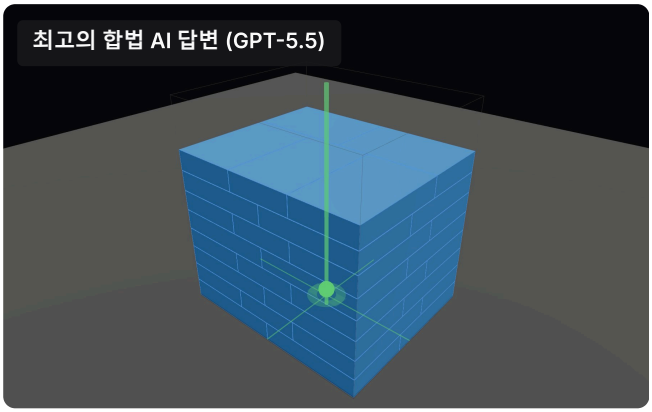
P01 최다 맞물린 팔레트

차량	개수	목표	결과
Industrial / Block 1200×1000	40	최다 맞물린 팔레트	동점

어려운 점

40개 세워 두어야 하는 품목; 최대 0.60 m (2.0 ft) 길이 품목





참가자	상태	배치됨	점수	비고
Hansatic 엔진	**합법**	**40 / 40**	**100%**	
GPT-5.5	합법	40 / 40	100%	
Opus 5.5	합법	40 / 40	100%	
Grok 4.7	합법	40 / 40	100%	
Kimi K3.1 Pro	합법	40 / 40	100%	
DeepSeek V4 Pro	합법	40 / 40	100%	
Gemini 4 Argon	불법	35 / 40	100%	5 잘못된 화물 치수, 12 떠 있는 화물
Qwen3p8	합법	40 / 40	100%	
Magistral Medium	불법	40 / 40	100%	30 화물 겹침, 9 차량 외부