Hansatic

तकनीकी श्वेतपत्र

न्यूरो-सिंबॉलिक लोड प्लानिंग

एक पैकिंग इंजन और आठ फ्रंटियर AI मॉडल वास्तविक ग्राहक लोड पर

नवंबर 2026

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

सहयोगी पृष्ठ, जिसमें हर प्लान का 3D दृश्य: hansatic.com/hi/neuro-symbolic-load-planning

सारांश

हम एक नियतात्मक पैकिंग इंजन को जनरेटिव AI के साथ जोड़ रहे हैं। संयोजन बनाने से पहले, हमने प्रत्येक भाग को अलग-अलग मापा।

- **डेटा:** 20 वास्तविक ग्राहक लोड-प्लानिंग मामले, 26 स्कोर किए गए रन। इनमें समुद्री कंटेनर, **कर्टेनसाइडर**, 53 फीट (16.2 मीटर) ड्राई वैन, **AKE** एयर कंटेनर और इंटरलॉक्ड पैलेट बिल्ड शामिल हैं।
- **प्रतिभागी:** हंसाटिक का उत्पादन पैकिंग इंजन और आठ मुख्यधारा के फ्रंटियर LLM। प्रत्येक मॉडल को एक JSON उत्तर में हर आइटम के लिए स्थिति और ओरिएंटेशन देना होता है, बिना किसी टूल के।
- **निर्णायक:** एक नियतात्मक वैलिडेटर हर प्रतिभागी को एक ही तरीके से स्कोर करता है। यह जाँचता है:
  - ज्यामिति;
  - 95% आधार समर्थन;
  - स्टैकिंग और नाजुक सीमाएँ;
  - ओरिएंटेशन फ्लैग;
  - पतली ढेरियाँ;
  - पेलोड और एक्सल लोड;
  - अनुदैर्ध्य **CoG** बैंड।

परिणाम:

- **इंजन** 26 में से 26 रन पर वैध है। प्रत्येक रन पर सर्वश्रेष्ठ वैध मॉडल उत्तर के मुकाबले, यह 18 पर आगे है (13 स्पष्ट रूप से, 5 मामूली अंतर से) और 8 पर बराबर। यह किसी पर पीछे नहीं है।
- **सबसे मजबूत मॉडल,** GPT-5.5, 26 में से 25 रन पर वैध है, जिसमें 253 टुकड़ों तक के मैनिफेस्ट शामिल हैं।
- **आठ मॉडलों में से चार** एक तिहाई से कम रन पर वैध हैं।
- **मॉडल किसमें अच्छे हैं।** सर्वश्रेष्ठ मॉडल उत्तरों ने दो लेआउट रणनीतियों का उपयोग किया जो इंजन के पास नहीं थीं। दोनों अब नियतात्मक इंजन रणनीतियाँ हैं।

हम निष्कर्ष निकालते हैं कि इंजन आज बेहतर प्लानर है, और मॉडल की रचनात्मकता को कैप्चर करना सार्थक है। इसके लिए हमारी वास्तुकला न्यूरो-सिंबॉलिक है:

- मॉडल रणनीतियाँ प्रस्तावित करते हैं;
- इंजन हर योजना बनाता है;
- वैलिडेटर केवल वैध सुधारों को रखता है।

हम मामले, वैलिडेटर, प्रॉम्प्ट और स्कोरिंग कोड जारी करते हैं।

1. बेंचमार्क डिज़ाइन

1.1 मामले

प्रत्येक मामला हंसाटिक ऐप से लिया गया वास्तविक ग्राहक लोड है, जिसे 1:1 कॉपी किया गया है: आयाम, वजन, मात्राएँ, फ्लैग और वाहन। आइटम आईडी बदले गए हैं और कार्गो नाम हटा दिए गए हैं।

सेट	मामले	रन	वाहन	टुकड़े
मानक (H01-H12)	12	16	एक 20' ड्राई, दो 40' ड्राई, चार 40' हाई क्यूब, एक 45' हाई क्यूब, चार 13.6 मीटर (44.6 फीट) कर्टेनसाइडर	14-83
बड़े (BV1-BV6)	6	8	तीन 40' हाई क्यूब, एक 40' ड्राई, एक 13.6 मीटर (44.6 फीट) कर्टेनसाइडर, एक 53 फीट (16.2 मीटर) ड्राई वैन	103-253
विशेष (U01, P01)	2	2	AKE (LD3) ढलान वाली दीवार के साथ; 1200 x 1000 मिमी (47 x 39 इंच) पैलेट	50 और 40

रन और उद्देश्य:

- **ट्रक** दो बार स्कोर किए जाते हैं: लोडिंग मीटर (LDM) और पैलेट पोजीशन पर। इससे 12 मानक मामलों से 16 रन और 6 बड़े मामलों से 8 रन प्राप्त होते हैं।
- **BV5 और BV6** पूरी तरह फिट नहीं होते। इनमें अधिक आइटम रखना मुख्य उद्देश्य है।
- **U01 और P01** में ज्ञात इष्टतम है:
  - U01: 50 क्यूब्स में से 24, क्योंकि ढलान वाली दीवार के कारण अधिक नहीं रखा जा सकता;
  - P01: 40 में से 40 बैग 100% इंटरलॉक के साथ।

मामलों में बाधाएँ:

- गैर-स्टैक करने योग्य और केवल ऊपर रखने वाले (नाजुक) आइटम;
- प्रति-आइटम स्टैक-वजन सीमाएँ;
- इस तरह ऊपर और झुकाव निषेध फ्लैग;
- लंबी छड़ें (BV1, कंटेनर की लंबाई तक);
- एक भारी आइटम का प्रभुत्व (H02);
- मिश्रित पैलेट और कार्टन लोड (BV3)।

1.2 सत्यापनकर्ता

सत्यापनकर्ता (v2.3.0) हर योजना की जाँच एक ही तरीके से करता है। एक योजना तभी वैध होती है जब हर हार्ड चेक पास हो जाए।

चेक	नियम
सीमाएँ	हर आइटम कार्गो स्पेस के अंदर हो, जिसमें AKE ढलान भी शामिल है
ओवरलैप	कोई दो आइटम एक ही आयतन साझा न करें

चेक	नियम
समर्थन	हर आइटम के आधार का कम से कम 95% फर्श या नीचे के आइटम पर टिका हो
स्टैकिंग	गैर-स्टैक करने योग्य आइटम कुछ नहीं उठाते; केवल ऊपर रखने वाले आइटम कुछ नहीं उठाते और सबसे ऊपर होते हैं; प्रति-आइटम स्टैक-वजन सीमाएँ लागू रहती हैं
ओरिएंटेशन	इस तरह ऊपर और झुकाव निषेध फ्लैग का पालन किया गया हो
पतली ढेरियाँ	मुक्त खड़ी स्टैक की ऊँचाई-से-आधार अनुपात सीमा
पेलोड	कुल वजन वाहन की सीमा के भीतर हो
एक्सल (ट्रक)	आगे और पीछे के एक्सल लोड सीमाओं के भीतर हों, आइटम के द्रव्यमान केंद्र से गणना की गई
CoG (कंटेनर)	अनुदैर्ध्य द्रव्यमान केंद्र अनुमत बैंड के अंदर हो
मल्टी-स्टॉप	स्टॉप के अनुसार अंतिम अंदर, पहला बाहर, जहाँ मामले में स्टॉप हों

उलटने का जोखिम गणना किया जाता है और रिपोर्ट किया जाता है, लेकिन स्कोर नहीं किया जाता।

1.3 रैंकिंग

प्रत्येक रन पर लागू:

1. एक वैध योजना अवैध योजना को हराती है।
2. अधिक आइटम रखने वाला जीतता है।
3. रन का उद्देश्य तय करता है:
    - कम लोड लंबाई (LDM);
    - या कम पैलेट पोजीशन;
    - या, P01 के लिए, अधिक इंटरलॉक शेयर, फिर कम ऊँचाई।
4. पैलेट पोजीशन पर टाई होने पर LDM से टूटता है।

एक काउंटर हर प्रतिभागी के लिए पैलेट पोजीशन गिनता है: हंसाटिक का उत्पादन काउंटर।

मार्जिन। हम जीत को कहते हैं:

- **स्पष्ट** यदि यह LDM पर कम से कम 2%, कम से कम एक पैलेट पोजीशन, या कम से कम एक अधिक आइटम रखा गया हो;
- **संकीर्ण** अन्यथा।

इंजन समय बजट के अंदर चलता है, इसलिए 2% से कम का एकल-रन अंतर रन-टू-रन भिन्नता हो सकता है (खंड 6.2)।

1.4 प्रतिभागी

मॉडल। आठ लैब्स के मुख्यधारा के फ्रंटियर मॉडल:

- 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)

सेटिंग्स:

- **रन:** प्रति मामला एक।
- **आउटपुट सीमा:** 64k टोकन, या प्रदाता की निचली अधिकतम (Gemini: 32k)।
- **बैचिंग:** प्रदाता बैच API जहाँ उपलब्ध हों।
- **उत्तर एकत्रित:** 2026-11-11।
- **कुल लागत:** लगभग $64।

इंजन। हंसाटिक का उत्पादन पैकर, 2026-11-20 का संस्करण (कमिट 9f66537c)। सेटिंग्स:

- उत्पादन विन्यास: 4 वर्कर और उत्पादन समय बजट, एक CPU कोर पर पिन किया गया;
- प्रति मामला एक रन, कोई बेस्ट-ऑफ नहीं;
- यह अपने 60% समर्थन नियम के अनुसार पैक करता है, और 95% पर हर मॉडल की तरह आंका जाता है।

1.5 प्रॉम्प्ट

- प्रति वाहन प्रकार एक प्रॉम्प्ट: कंटेनर, ट्रक, ULD और पैलेट।
- प्रत्येक प्रॉम्प्ट उस प्रकार पर लागू होने वाले सटीक नियम, उद्देश्य और आउटपुट स्कीमा (आइटम आईडी, स्थिति, ओरिएंटेशन) बताता है।
- हल किए गए लेआउट उदाहरण के रूप में नहीं दिए जाते।

पूर्ण टेम्प्लेट परिशिष्ट A में हैं।

2. परिणाम

2.1 प्रति मॉडल, सभी 26 रन

मॉडल	वैध	इंजन के बराबर	इंजन से पीछे
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

कोई भी मॉडल किसी भी रन पर इंजन को नहीं हरा सका।

मानक सेट (16 रन)।

मॉडल	वैध	बराबर	पीछे
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

बड़े सेट (8 रन)। किसी भी मॉडल ने बड़े रन पर इंजन के बराबर नहीं किया।

मॉडल	वैध
GPT-5.5	8
Grok 4.7	8
Opus 5.5	5

मॉडल	वैध
Kimi K3.1 Pro	2
DeepSeek V4 Pro	1
Magistral Medium	1
Gemini 4 Argon	0
Qwen3p8	0

विशेष सेट (2 रन)।

मॉडल	वैध	इष्टतम तक पहुँच
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 इंजन का सर्वश्रेष्ठ वैध मॉडल उत्तर के विरुद्ध प्रदर्शन

सेट	स्पष्ट रूप से आगे	मामूली रूप से आगे	बराबर	पीछे
मानक (16)	7	3	6	0
बड़ा (8)	6	2	0	0
विशेष (2)	0	0	2	0
सभी (26)	**13**	**5**	**8**	**0**

मानक सेट। प्रत्येक सेल में रखे गए आइटम, फिर लोड लंबाई या पैलेट पोजीशन दी गई है।

रन	इंजन	सर्वश्रेष्ठ वैध मॉडल	अंतर
H01	26/27, 11.69 मी (38.4 फीट)	Opus 26, 11.74 मी (38.5 फीट)	मामूली (0.4%)
H02	24/29, 5.78 मी (19.0 फीट)	GPT 18, 5.76 मी (18.9 फीट)	स्पष्ट (+6 रखे गए)
H03	31/32, 10.51 मी (34.5 फीट)	कोई वैध नहीं	स्पष्ट
H04	39, 10.44 मी (34.3 फीट)	Opus 39, 10.44 मी (34.3 फीट)	बराबर

रन	इंजन	सर्वश्रेष्ठ वैध मॉडल	अंतर
H05	49, 12.40 मी (40.7 फीट)	Opus 49, 12.40 मी (40.7 फीट)	बराबर
H06	83, 10.65 मी (35.0 फीट)	Grok 83, 11.41 मी (37.4 फीट)	स्पष्ट (6.6%)
H07	45, 9.60 मी (31.5 फीट)	GPT 45, 9.60 मी (31.5 फीट)	बराबर
H08	52, 5.92 मी (19.4 फीट)	Kimi 52, 6.12 मी (20.1 फीट)	स्पष्ट (3.3%)
H09 LDM	29/30, 13.36 मी (43.8 फीट)	GPT 29, 13.56 मी (44.5 फीट)	मामूली (1.5%)
H09 पैलेट	29/30, 29 पोजीशन	GPT 29, 29 पोजीशन	मामूली (LDM टाईब्रेकर)
H10 LDM	26, 12.20 मी (40.0 फीट)	DeepSeek, समान	बराबर
H10 पैलेट	26, 31 पोजीशन	DeepSeek, समान	बराबर
H11 LDM	23, 12.00 मी (39.4 फीट)	Opus 23, 12.20 मी (40.0 फीट)	मामूली (1.6%)
H11 पैलेट	23, 30 पोजीशन	Kimi 23, 32 पोजीशन	स्पष्ट (2 पोजीशन)
H12 LDM	13/14, 11.00 मी (36.1 फीट)	Gemini 13, 11.00 मी (36.1 फीट)	बराबर
H12 पैलेट	13/14, 25 पोजीशन	Kimi 13, 28 पोजीशन	स्पष्ट (3 पोजीशन)

बड़ा सेट।

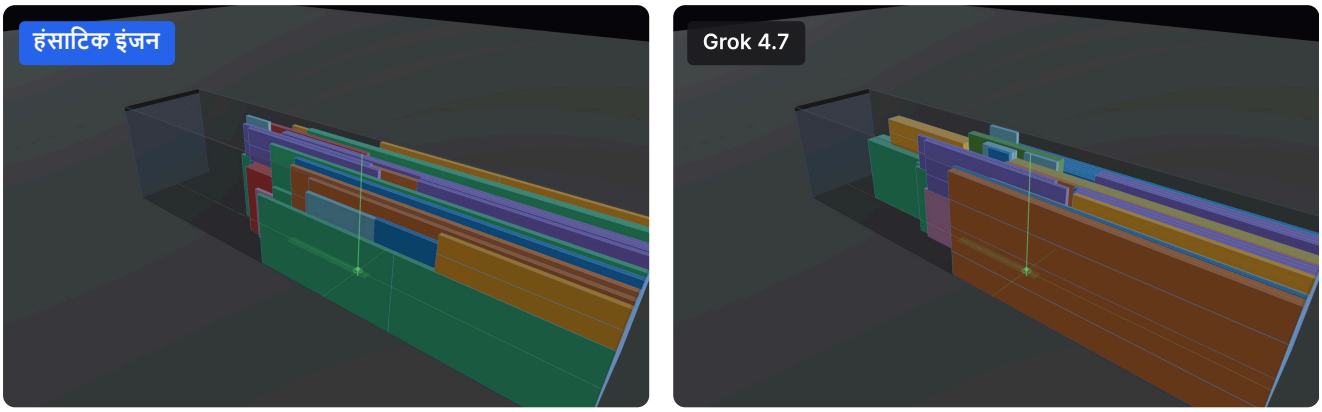
रन	इंजन	सर्वश्रेष्ठ वैध मॉडल	अंतर
BV1 (40' HC, 103 बार)	103, 9.17 मी (30.1 फीट)	Grok 103, 10.66 मी (35.0 फीट)	स्पष्ट (14.0%)
BV2 LDM (कर्टेनसाइडर, 153)	153, 11.52 मी (37.8 फीट)	Opus 153, 11.78 मी (38.6 फीट)	स्पष्ट (2.2%)
BV2 पैलेट	153, 29 पोजीशन	Opus 153, 29 पोजीशन	मामूली (LDM टाईब्रेकर)
BV3 LDM (53 फीट वैन, 119)	119, 6.83 मी (22.4 फीट)	Opus 119, 6.99 मी (22.9 फीट)	स्पष्ट (2.2%)
BV3 पैलेट	119, 17 पोजीशन	Opus 119, 18 पोजीशन	स्पष्ट (1 पोजीशन)
BV4 (145, केवल ऊपर नाजुक)	145, 8.57 मी (28.1 फीट)	Grok 145, 8.85 मी (29.0 फीट)	स्पष्ट (3.1%)
BV5 (136, केवल सीधा)	127, 12.00 मी (39.4 फीट)	Opus 124, 11.77 मी (38.6 फीट)	स्पष्ट (+3 रखे गए)
BV6 (253, 95% आयतन)	244, 11.98 मी (39.3 फीट)	GPT 238, 11.97 मी (39.3 फीट)	स्पष्ट (+6 रखे गए)

विशेष सेट।

रन	इंजन	मॉडल
U01 (AKE)	24/50, इष्टतम	Opus, Grok, Kimi, Qwen और Gemini भी 24 तक पहुँचे; GPT और DeepSeek 18 तक पहुँचे
P01 (पैलेट)	40/40, 100% इंटरलॉक, इष्टतम	हर वैध मॉडल उत्तर 40/40 इष्टतम पर पहुँचा

2.3 वास्तविक लोड पर अंतर का अर्थ

- **BV1:** 1.49 मी (4.9 फीट) छोटा, लगभग 40' कंटेनर की लंबाई का 12%।
- **H12 पैलेट:** 3 कम पैलेट पोजीशन, लगभग 13.6 मी (44.6 फीट) कर्टेनसाइडर डेक का 9%।
- **BV6:** 6 अधिक टुकड़े रखे गए जो पूरी तरह फिट नहीं होते, इसलिए दूसरे वाहन के लिए 6 कम टुकड़े।



BV1: हंसाटिक इंजन 9.17 मी (30.1 फीट), बाएँ; सर्वश्रेष्ठ वैध मॉडल उत्तर (Grok 4.7) 10.66 मी (35.0 फीट), दाएँ।

2.4 लागत और समय

	इंजन	मॉडल
प्रति योजना समय	नगण्य	तुलना नहीं की गई: बैच API अनुरोधों को कतारबद्ध करते हैं
प्रति योजना लागत	नगण्य	$0.10-0.80

3. विफलता विश्लेषण

208 मॉडल रन (8 मॉडल × 26) में से 95 योजनाएँ वैध हैं। अन्य 113 तीन अलग-अलग तरीकों से विफल होती हैं।

मॉडल	वैध	कोई उत्तर नहीं (आउटपुट सीमा)	प्रदाता त्रुटि	नियमों का उल्लंघन
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
सभी	**95**	**65**	**4**	**44**

3.1 कोई उत्तर नहीं: मॉडल सोचने की जगह खत्म कर देता है

65 रन में, मॉडल ने तर्क किया जब तक कि वह आउटपुट सीमा तक नहीं पहुँच गया और कभी योजना नहीं लिखी। एक और उत्तर (Magistral, H03) सामान्य रूप से समाप्त हुआ लेकिन वैध JSON नहीं था; इसे यहाँ गिना गया है।

- यह सबसे आम विफलता है, और यह आठ मॉडलों में से छह को प्रभावित करती है।
- Kimi, DeepSeek और Qwen इस तरह अपने अधिकांश रन खो देते हैं। उनकी वैध योजनाएँ ज्यादातर छोटे मामलों पर होती हैं।
- Opus इस तरह 6 रन खो देता है, जिनमें BV1, BV4 और BV6 शामिल हैं, जो तीन सबसे बड़े मैनिफेस्ट हैं।

उत्पादन में योजनाकार के लिए, कोई उत्तर वैध उत्तर न होने जैसी ही विफलता है: डॉक को अभी भी एक योजना की आवश्यकता होती है।

3.2 प्रदाता त्रुटियाँ

Grok ने 4 रन पर HTTP 403 लौटाया, क्योंकि प्रदाता की खर्च सीमा पूरी हो गई थी। इन रनों को दोबारा नहीं भेजा गया और इन्हें विफलता के रूप में गिना गया।

3.3 नियमों का उल्लंघन करने वाली योजनाएँ

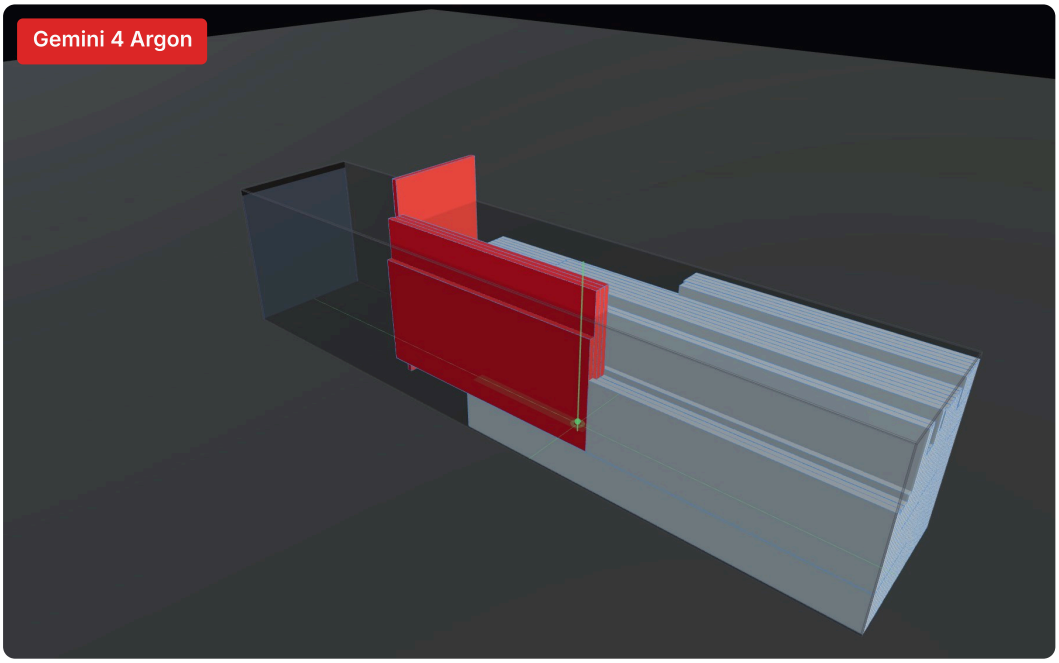
44 उत्तर पूर्ण योजनाएँ हैं जो कम से कम एक कठोर नियम का उल्लंघन करती हैं। लगभग सभी दो मॉडलों से आती हैं: Gemini (19) और Magistral (23)। Grok के 2 हैं।

इन 44 योजनाओं में से प्रत्येक प्रकार के उल्लंघन वाली योजनाओं की संख्या:

टूटा हुआ नियम	योजनाएँ
माल स्थान के बाहर वस्तु	27
ओवरलैपिंग वस्तुएँ	27
तैरती वस्तुएँ (95% समर्थन से कम)	21
वस्तु आयाम जो किसी अनुमत ओरिएंटेशन से मेल नहीं खाते	17
गैर-स्टैक करने योग्य या नाजुक वस्तुओं पर भार, या स्टैक सीमा से अधिक	10
छत से ऊपर स्टैक	9
गुरुत्वाकर्षण केंद्र (CoG) बैंड के बाहर	7
एक्सल ओवरलोड	4
नो-टिल्ट वस्तु झुकी हुई	1
AKE कंटूर के बाहर	1

Grok की दो अवैध योजनाएँ प्रत्येक एक नियम का उल्लंघन करती हैं: एक रन पर कुचली हुई वस्तु, दूसरे पर गुरुत्वाकर्षण केंद्र।

ये योजनाएँ 3D दृश्य में विश्वसनीय लगती हैं। दोष अक्सर कुछ वस्तुएँ होती हैं, जो सेंटीमीटर से तैरती या ओवरलैप होती हैं। केवल पूर्ण ज्यामितीय जाँच से ये पकड़ में आती हैं।



BV1, Gemini 4 Argon: 103 में से 77 वस्तुएँ रखी गईं। हाइलाइट की गईं: वे वस्तुएँ जो तैरती हैं, ओवरलैप होती हैं या माल स्थान से बाहर जाती हैं।

3.4 वैध, लेकिन पीछे

- **अधूरी।** BV5 और BV6 पर, सर्वश्रेष्ठ वैध मॉडल उत्तर इंजन से 3 और 6 कम वस्तुएँ रखते हैं।

- **लंबी।** उन रनों पर जहाँ हर वस्तु फिट होती है, सर्वश्रेष्ठ वैध मॉडल उत्तर 24 ट्रक और कंटेनर रनों में से 13 पर इंजन से लंबा होता है। सबसे बड़े अंतर:
  - BV1 पर 1.49 मीटर (4.9 फीट);
  - H06 पर 0.76 मीटर (2.5 फीट)।

3.5 मॉडल क्या अच्छा करते हैं

सबसे मजबूत उत्तर उन रणनीतियों का उपयोग करते हैं जो एक अनुभवी टीम इस्तेमाल करती है:

- **भारी क्रेट सामने की दीवार के पास।**
- **गैर-स्टैक करने योग्य वस्तुओं के ऊपर ऊँचाई का उपयोग।** BV3 पर, Opus कार्टन को उनकी तरफ 559 मिमी (22 इंच) ऊँची फर्श के रूप में बिछाता है और उस पर 19 गैर-स्टैक करने योग्य उपकरण खड़े करता है।

इंजन के पास इनमें से कोई भी रणनीति नहीं थी। दोनों अब नियतात्मक इंजन रणनीतियाँ हैं। यह वही रचनात्मकता है जिसे न्यूरो-सिंबोलिक लूप कैप्चर करने के लिए डिज़ाइन किया गया है (खंड 5)।

4. उपकरण भुजा

दूसरी भुजा में, मॉडल एक उपकरण `pack_cargo` के माध्यम से इंजन को कॉल कर सकता है और फिर किसी भी लोडिंग योजना के साथ उत्तर दे सकता है।

अधिकांश मॉडल मैनिफेस्ट को उपकरण कॉल में कॉपी करते हैं और इंजन की योजना को बिना बदलाव के लौटाते हैं। विफलताएँ कॉपी करने की त्रुटियों के कारण होती हैं:

- मैजिस्ट्रल ने 100 वस्तुओं को गलत आयाम या फ्लैग के साथ भेजा;
- किमी ने H10 पर गलत कॉल किया;
- क्केन ने 4 रन पर विफल रहा।

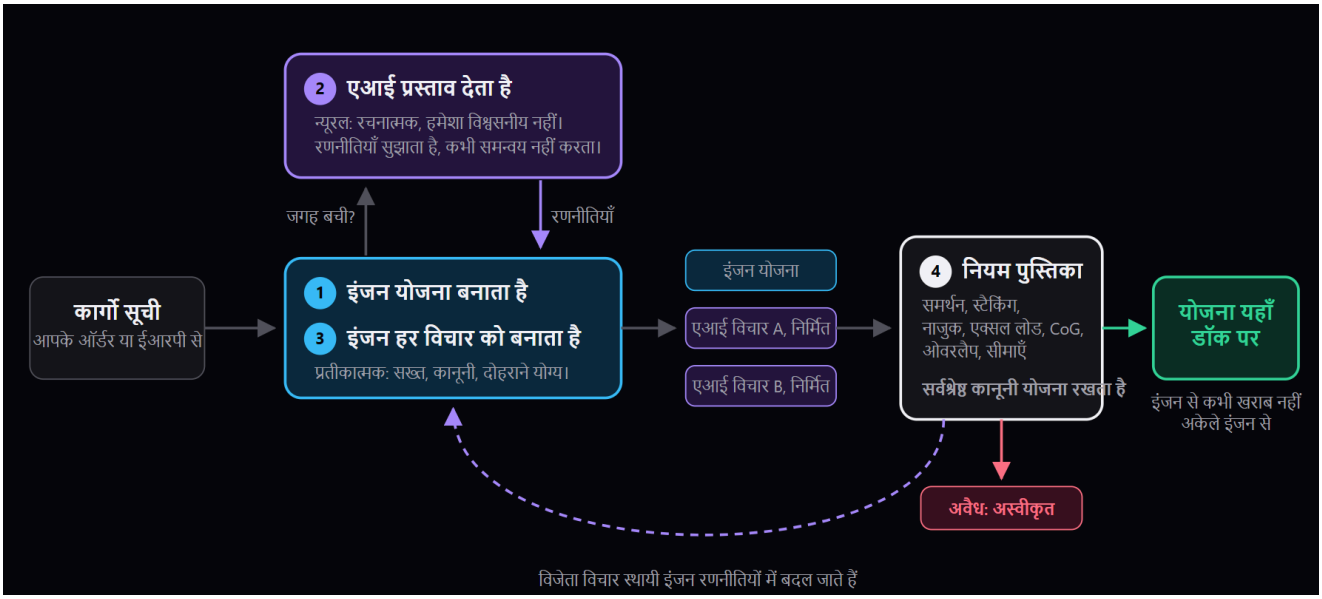
निहितार्थ: किसी भी संयुक्त प्रणाली में, इंजन को मैनिफेस्ट सीधे स्रोत प्रणाली से पढ़ना चाहिए, मॉडल के माध्यम से कभी नहीं।

5. न्यूरो-सांकेतिक वास्तुकला

5.1 दो प्रकार की बुद्धिमत्ता

	पैकिंग इंजन (सांकेतिक)	जनरेटिव एआई (न्यूरल)
नियम	हर एन्कोडेड नियम का हर बार पालन करता है	दिए गए नियमों का उल्लंघन करता है
पुनरावृत्ति	एक ही लोड, एक ही योजना	एक ही लोड के लिए अलग योजना मिल सकती है
ऑडिट	हर प्लेसमेंट का कारण होता है	ऑडिट करना कठिन
सीमा	केवल उन रणनीतियों तक सीमित जो इसे पता हैं	नई रणनीतियाँ खोजता है
स्केल	253 टुकड़ों को संभालता है	कई मॉडल बड़े मैनिफेस्ट पर कोई उत्तर नहीं देते

5.2 लूप



1. **आधार रेखा।** इंजन मैनिफेस्ट से योजना बनाता है, जो सीधे स्रोत प्रणाली से पढ़ा जाता है।
2. **ट्रिगर।** मॉडल को केवल तब बुलाया जाता है जब आधार रेखा में जगह बचती है: अनप्लेस्ड आइटम, कम भराव, या क्लीयरेंस लक्ष्य छूट जाता है।
3. **प्रस्ताव।** मॉडल रणनीति पैरामीटर लौटाता है, कभी निर्देशांक नहीं:
   - जोन;
   - क्रम;
   - ओरिएंटेशन प्राथमिकताएँ;
   - फर्श और कैप संरचनाएँ।
4. **निष्पादन।** इंजन हर प्रस्ताव के लिए एक योजना बनाता है।
5. **निर्णय।** वैलिडेटर सर्वश्रेष्ठ वैध योजना रखता है। आधार रेखा हमेशा एक उम्मीदवार होती है।

6. **संक्षेपण।** बार-बार जीतने वाला प्रस्ताव स्थायी इंजन रणनीति बन जाता है, जैसे फ्रंट-वॉल और कार्टन-फ्लोर रणनीतियाँ बनीं।

5.3 निर्माण द्वारा गुण

- **इंजन अकेले से कभी खराब नहीं**, क्योंकि आधार रेखा हमेशा एक उम्मीदवार होती है।
- **कोई अवैध योजना डॉक तक नहीं पहुँचती**, क्योंकि वैलिडेटर हर उम्मीदवार का निर्णय करता है।
- **कोई ट्रांसक्रिप्शन जोखिम नहीं**, क्योंकि मॉडल कभी मैनिफेस्ट या निर्देशांकों को छूता नहीं (खंड 4)।
- **मॉडल के आउटपुट सीमा से कोई विफलता नहीं**, क्योंकि प्रस्ताव कुछ पैरामीटर होते हैं, 253-आइटम योजना नहीं (खंड 3.1)।
- **सीमित एआई लागत**, क्योंकि मॉडल केवल ट्रिगर पर चलता है।

हम इस लूप का निर्माण कर रहे हैं। अगले संस्करण में इसे अपने प्रवेशक के रूप में स्कोर किया जाएगा, जिसमें प्रति योजना लागत और समय शामिल होगा।

6. दायरा और सीमाएँ

6.1 वैधता का दायरा

कानूनी का अर्थ है v2.3.0 में एन्कोड किए गए नियमों के तहत कानूनी। वैधक मल्टी-स्टॉप लास्ट इन, फर्स्ट आउट से आगे की लोडिंग क्रम की जाँच नहीं करता।

6.2 विचलन

प्रत्येक प्रतिभागी प्रति केस एक बार चलता है।

- **इंजन** समय बजट के अंदर रणनीतियों की दौड़ लगाता है, इसलिए मशीन लोड बदल सकता है कि कौन सी रणनीति पहले पूरी होती है। H02 पर हमने 5.63-5.79 मीटर (18.5-19.0 फीट) के बीच मापा।
- **मॉडल** का एक बार नमूना लिया जाता है।

2% से कम के एकल-रन अंतर को संकीर्ण बताया गया है।

6.3 मॉडल स्तर और आउटपुट सीमा

- हमने प्रत्येक लैब के मुख्यधारा के फ्रंटियर मॉडल को चलाया। प्रीमियम स्तर (Anthropic Fable, OpenAI GPT-6 Astra) नहीं चलाए गए।
- एक बड़ी आउटपुट सीमा कुछ नो-आंसर रन (खंड 3.1) को प्लान में बदल सकती थी।

6.4 केस स्रोत

सभी केस एक प्लेटफॉर्म के उपयोगकर्ताओं से आए हैं। अन्य शिपर्स का कार्गो मिश्रण भिन्न हो सकता है।

7. अगले कदम

7.1 प्रैक्टिशनर नियम

अगले दो प्रैक्टिशनर नियम इंजन और वैलिडेटर v3 में जोड़े जाएंगे:

- **छत की निकासी:** छत के नीचे कम से कम 50 मिमी (2.0 इंच) या 25 किलोग्राम (55 पाउंड) से अधिक के स्टैक्ड आइटम के ऊपर 100 मिमी (3.9 इंच)।
- प्रत्येक फोर्कलिफ्ट पंक्ति के लिए **साइड प्ले**।

हमने हर प्रतिभागी को छत के नियम के साथ पुनः स्कोर किया। फर्श पर रखे आइटम जो झुक नहीं सकते और छत से 50 मिमी (2.0 इंच) कम ऊंचाई वाले हैं, उन्हें छूट दी गई है, क्योंकि मामला उन्हें मजबूर करता है।

- **इंजन** 16 मानक रन में से 9 और 8 बड़े रन में से 1 पर नियम बनाए रखता है।
- **6 रन पर, एक मॉडल उत्तर नियम बनाए रखता है** इंजन के प्लान के करीब। ये इंजन के पहले लक्ष्य हैं:
  - H02 (GPT);
  - H06 (Grok, 11.41 मी / 37.4 फीट);

- H12 लोडिंग मीटर (Gemini, 11.00 मी / 36.1 फीट);
- H12 पैलेट पोजीशन (Kimi, 28 पोजीशन);
- BV3 लोडिंग मीटर (GPT, 9.14 मी / 30.0 फीट);
- BV5 (Opus, 124 रखे गए)।

- **8 रन पर, कोई भी प्रतिभागी नियम नहीं बनाए रखता:** H01, H03, H08, BV1, दोनों BV2 रन, BV3 पैलेट, BV6।

7.2 अगला संस्करण

- वैलिडेटर v3, जिसमें छत की निकासी, साइड प्ले और बिल्ड-सीक्वेंस चेक शामिल होगा, हर प्रतिभागी पर लागू किया जाएगा।
- नए वास्तविक ग्राहक मामले।
- प्रत्येक लैब के सबसे सक्षम मॉडल, प्रीमियम टियर सहित, प्रति मॉडल कई रन और भिन्नता रिपोर्ट की जाएगी।
- न्यूरो-सिंबॉलिक लूप को एक प्रतिभागी के रूप में शामिल किया जाएगा।

8. रिलीज़

प्रकाशित:

- प्रत्येक मामले का पूरा विवरण;
- सत्यापनकर्ता;
- प्रॉम्प्ट;
- स्कोरिंग और रैंकिंग कोड;
- प्रत्येक रन के लिए: भेजा गया अनुरोध, प्रदाता का अपरिवर्तित उत्तर और पार्स किया गया प्लान।

दर्शक: प्रत्येक प्रतिभागी का प्रत्येक प्लान, 3D में अपने प्रॉम्प्ट और कच्चे उत्तर के साथ: hansatic.com/hi/neuro-symbolic-load-planning.

बंद:

- पैलेट-पोज़िशन काउंटर (Hansatic का उत्पादन काउंटर, सभी प्रतिभागियों पर समान रूप से लागू);
- इंजन;
- स्रोत प्लान की आईडी।

```
```

परिशिष्ट A. प्रॉम्प्ट टेम्प्लेट

प्रत्येक वाहन प्रकार के लिए एक सिस्टम प्रॉम्प्ट, जैसा भेजा गया। उपयोगकर्ता संदेश उद्देश्य बताता है, फिर केस को JSON प्रारूप में प्रस्तुत करता है; यहाँ आइटम सूची को संक्षिप्त किया गया है। प्रॉम्प्ट अंग्रेजी में दिखाए गए हैं, वही भाषा जिसमें मॉडल्स को भेजे गए थे।

समुद्री कंटेनर (BV1, 40' हाई क्यूब)

सिस्टम प्रॉम्प्ट

```
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"]}
```

उपयोगकर्ता संदेश (आइटम सूची संक्षिप्त की गई)

```
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"
 ]
}
```


ट्रक (H09, 13.6 मीटर (44.6 फीट) पर्दानुमा)

सिस्टम प्रॉम्प्ट

```
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"]}
```

उपयोगकर्ता संदेश (आइटम सूची संक्षिप्त की गई)

```
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"
```

```
 ]
}
```

वायु कंटेनर (U01, AKE जिसमें ढलान वाली दीवार है)

सिस्टम प्रॉम्प्ट

```
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"]}
```

उपयोगकर्ता संदेश (आइटम सूची संक्षिप्त की गई)

```
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"
 ]
}
```


पैलेट (P01, 1200 × 1000 मिमी (47 x 39 इंच))

सिस्टम प्रॉम्प्ट

```
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"]}
```

उपयोगकर्ता संदेश (आइटम सूची संक्षिप्त की गई)

```
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"
```

```
  ]
}
```

परिशिष्ट B. सभी 26 रन

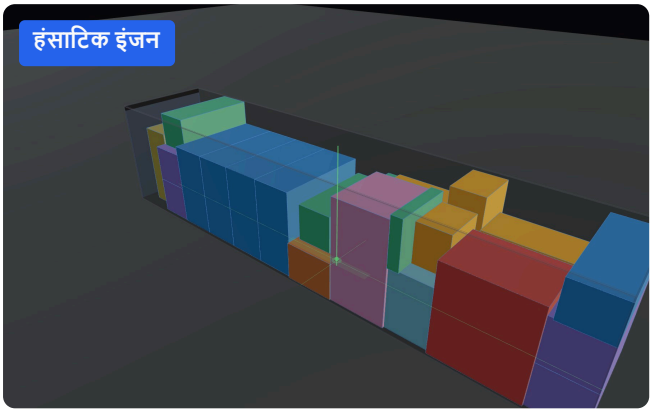
प्रत्येक रन के लिए एक पृष्ठ: लोड, इंजन की योजना के बगल में सर्वश्रेष्ठ वैध मॉडल उत्तर, और प्रत्येक प्रतिभागी का परिणाम।

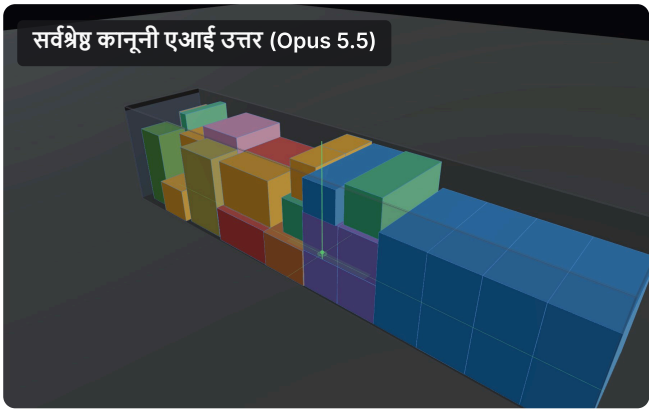
H01 सबसे छोटा लोड

वाहन	टुकड़े	उद्देश्य	परिणाम
40' High Cube	27	सबसे छोटा लोड	इंजन मामूली रूप से आगे

जो इसे कठिन बनाता है

27 आइटम जो सीधे खड़े रहने चाहिए





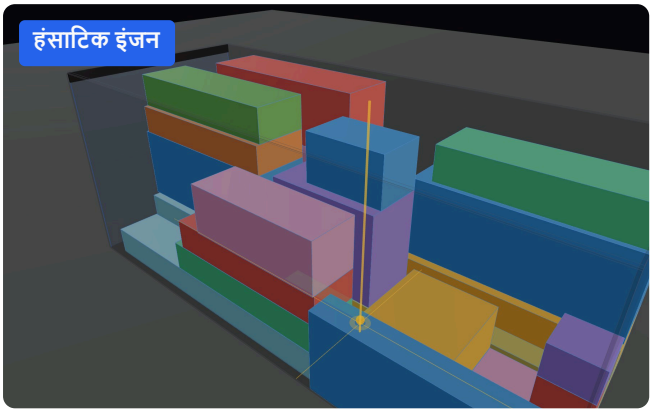
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**26 / 27**	**11.69 m (38.4 ft)**	
GPT-5.5	**कानूनी**	25 / 27	11.88 m (39.0 ft)	
Opus 5.5	**कानूनी**	26 / 27	11.74 m (38.5 ft)	
Grok 4.7	**कानूनी**	25 / 27	11.37 m (37.3 ft)	
Kimi K3.1 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
DeepSeek V4 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Gemini 4 Argon	**अवैध**	22 / 27	11.56 m (37.9 ft)	1 गलत आइटम आकार, 1 वाहन के बाहर, 1 तैरते हुए आइटम
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	**अवैध**	27 / 27	12.09 m (39.7 ft)	7 आइटम ओवरलैप हो रहे हैं, 14 वाहन के बाहर, 3 ऊंचाई सीमा से अधिक, 1 तैरते हुए आइटम, 1 गुरुत्वाकर्षण केंद्र (CoG) गलत

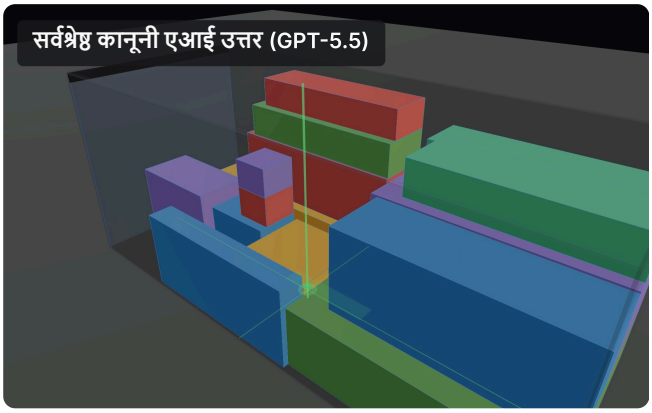
H02 सबसे छोटा लोड

वाहन	टुकड़े	उद्देश्य	परिणाम
20' Dry Container	29	सबसे छोटा लोड	इंजन स्पष्ट रूप से आगे

जो इसे कठिन बनाता है

29 आइटम जो सीधे खड़े रहने चाहिए; 5.00 m (16.4 ft) तक लंबे आइटम; सबसे भारी आइटम 1,550 kg (3,417 lb)





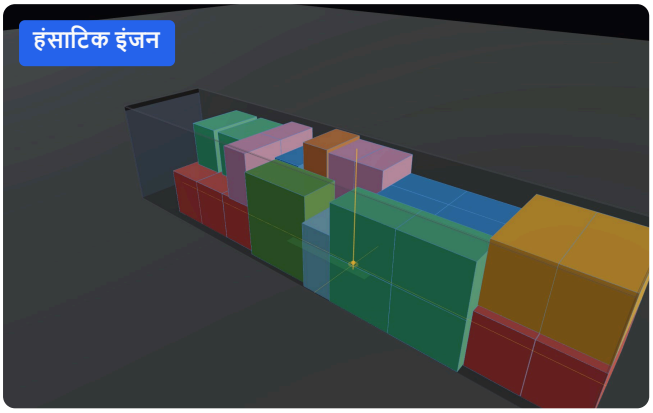
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**24 / 29**	**5.78 m (19.0 ft)**	
GPT-5.5	**कानूनी**	18 / 29	5.76 m (18.9 ft)	
Opus 5.5	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Grok 4.7	प्रदाता त्रुटि	-	-	प्रदाता ने अनुरोध अस्वीकार कर दिया।
Kimi K3.1 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
DeepSeek V4 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Gemini 4 Argon	**अवैध**	23 / 29	5.00 m (16.4 ft)	6 गलत आइटम आकार, 13 आइटम ओवरलैप हो रहे हैं, 9 तैरते हुए आइटम, 9 दबे हुए आइटम
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	**अवैध**	29 / 29	7.97 m (26.1 ft)	37 आइटम ओवरलैप हो रहे हैं, 23 वाहन के बाहर, 1 ऊंचाई सीमा से अधिक, 7 तैरते हुए आइटम, 5 दबे हुए आइटम, 2 गुरुत्वाकर्षण केंद्र (CoG) गलत

H03 सबसे छोटा लोड

वाहन	टुकड़े	उद्देश्य	परिणाम
40' High Cube	32	सबसे छोटा लोड	इंजन स्पष्ट रूप से आगे

जो इसे कठिन बनाता है

25 आइटम जो सीधे खड़े रहने चाहिए; सबसे भारी आइटम 2,680 kg (5,908 lb)



इस रन पर किसी भी एआई मॉडल ने कानूनी योजना नहीं दी।

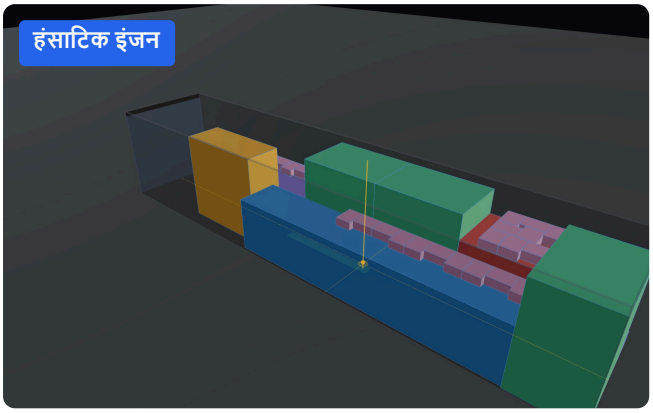
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**31 / 32**	**10.51 m (34.5 ft)**	
GPT-5.5	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Opus 5.5	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Grok 4.7	**अवैध**	31 / 32	12.03 m (39.5 ft)	1 दबे हुए आइटम
Kimi K3.1 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
DeepSeek V4 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Gemini 4 Argon	**अवैध**	31 / 32	11.81 m (38.7 ft)	1 गलत आइटम आकार, 12 आइटम ओवरलैप हो रहे हैं, 6 वाहन के बाहर, 3 ऊंचाई सीमा से अधिक, 2 तैरते हुए आइटम, 3 दबे हुए आइटम
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	कोई योजना नहीं दी गई	-	-	उत्तर में कोई वैध योजना नहीं थी।

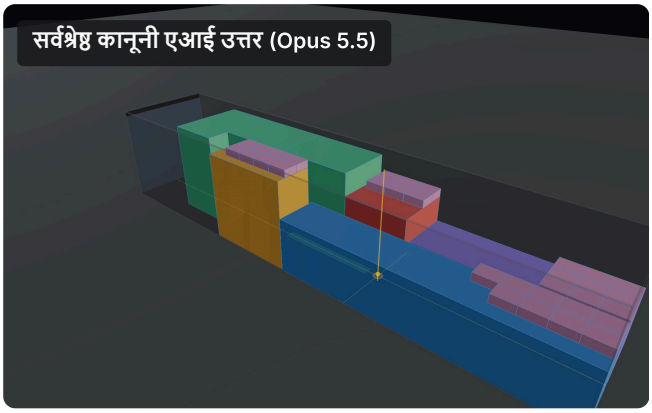
H04 सबसे छोटा लोड

वाहन	टुकड़े	उद्देश्य	परिणाम
40' Dry Container	39	सबसे छोटा लोड	बराबरी

जो इसे कठिन बनाता है

6.30 m (20.7 ft) तक लंबे आइटम; सबसे भारी आइटम 998 kg (2,200 lb)





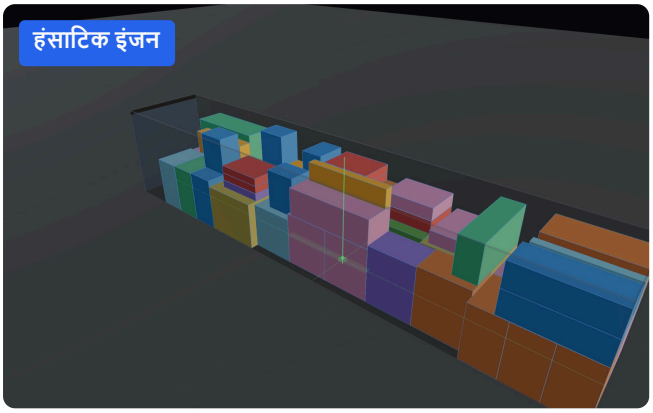
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**39 / 39**	**10.44 m (34.3 ft)**	
GPT-5.5	**कानूनी**	39 / 39	10.72 m (35.2 ft)	
Opus 5.5	**कानूनी**	39 / 39	10.44 m (34.3 ft)	
Grok 4.7	**कानूनी**	39 / 39	11.10 m (36.4 ft)	
Kimi K3.1 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
DeepSeek V4 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Gemini 4 Argon	**अवैध**	29 / 39	6.30 m (20.7 ft)	10 गलत आइटम आकार, 6 तैरते हुए आइटम
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	**अवैध**	38 / 39	10.20 m (33.5 ft)	1 गलत आइटम आकार, 27 आइटम ओवरलैप हो रहे हैं, 6 वाहन के बाहर, 24 तैरते हुए आइटम, 1 दबे हुए आइटम, 1 गुरुत्वाकर्षण केंद्र (CoG) गलत

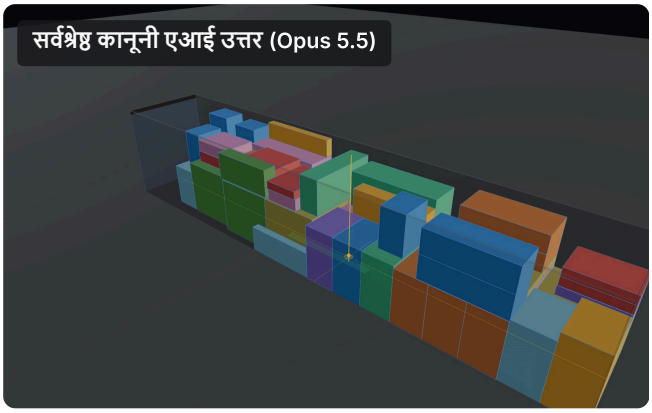
H05 सबसे छोटा लोड

वाहन	टुकड़े	उद्देश्य	परिणाम
45' High Cube	49	सबसे छोटा लोड	बराबरी

जो इसे कठिन बनाता है

49 आइटम जो सीधे खड़े रहने चाहिए; 27 आइटम जिनकी ढेर-भार सीमा है; सबसे भारी आइटम 814 kg (1,795 lb)





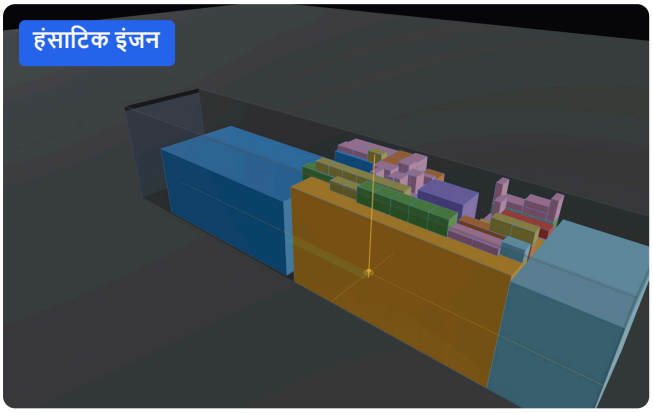
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**49 / 49**	**12.40 m (40.7 ft)**	
GPT-5.5	**कानूनी**	45 / 49	13.28 m (43.6 ft)	
Opus 5.5	**कानूनी**	49 / 49	12.40 m (40.7 ft)	
Grok 4.7	**कानूनी**	43 / 49	12.96 m (42.5 ft)	
Kimi K3.1 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
DeepSeek V4 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Gemini 4 Argon	**अवैध**	37 / 49	10.49 m (34.4 ft)	11 गलत आइटम आकार, 2 आइटम ओवरलैप हो रहे हैं, 5 वाहन के बाहर, 2 ऊंचाई सीमा से अधिक, 3 तैरते हुए आइटम, 2 दबे हुए आइटम
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	**अवैध**	47 / 49	12.73 m (41.8 ft)	2 गलत आइटम आकार, 31 आइटम ओवरलैप हो रहे हैं, 9 वाहन के बाहर, 14 ऊंचाई सीमा से अधिक, 21 तैरते हुए आइटम, 2 दबे हुए आइटम

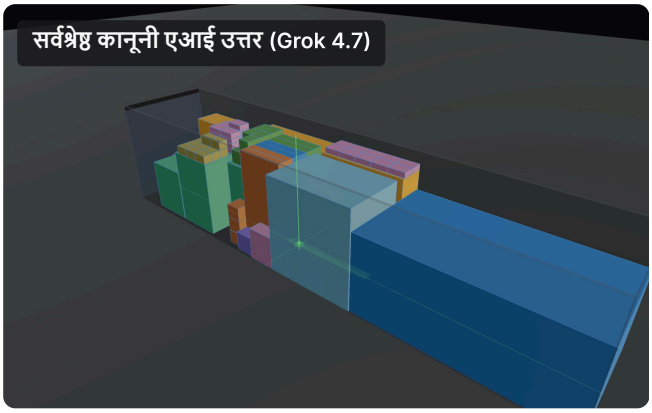
H06 सबसे छोटा लोड

वाहन	टुकड़े	उद्देश्य	परिणाम
40' High Cube	83	सबसे छोटा लोड	इंजन स्पष्ट रूप से आगे

जो इसे कठिन बनाता है

5 गैर-ढेर करने योग्य आइटम; 1 आइटम जो सीधे खड़े रहने चाहिए; सबसे भारी आइटम 3,020 kg (6,658 lb)





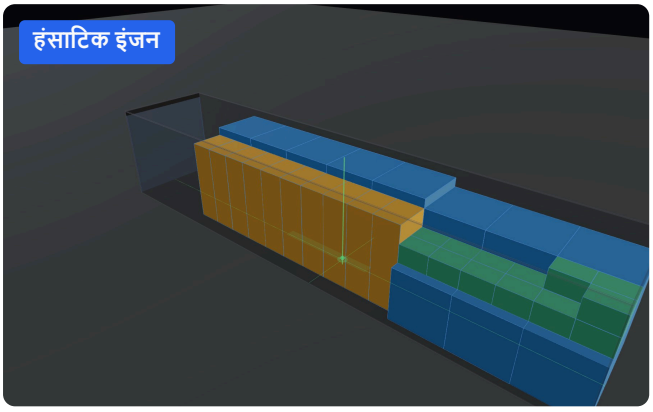
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**83 / 83**	**10.65 m (35.0 ft)**	
GPT-5.5	**कानूनी**	83 / 83	11.78 m (38.7 ft)	
Opus 5.5	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Grok 4.7	**कानूनी**	83 / 83	11.41 m (37.4 ft)	
Kimi K3.1 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
DeepSeek V4 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Gemini 4 Argon	**अवैध**	7 / 83	6.89 m (22.6 ft)	2 गलत आइटम आकार, 1 गुरुत्वाकर्षण केंद्र (CoG) गलत
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	**अवैध**	20 / 83	12.25 m (40.2 ft)	2 आइटम ओवरलैप हो रहे हैं, 4 वाहन के बाहर, 1 ऊंचाई सीमा से अधिक, 14 तैरते हुए आइटम, 1 गुरुत्वाकर्षण केंद्र (CoG) गलत

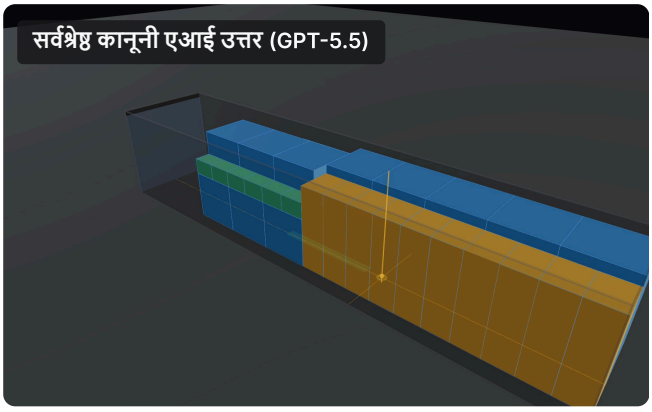
H07 सबसे छोटा लोड

वाहन	टुकड़े	उद्देश्य	परिणाम
40' Dry Container	45	सबसे छोटा लोड	बराबरी

जो इसे कठिन बनाता है

10 गैर-ढेर करने योग्य आइटम; 15 आइटम जो सीधे खड़े रहने चाहिए; 15 आइटम जिनकी ढेर-भार सीमा है; सबसे भारी आइटम 1,200 kg (2,646 lb)





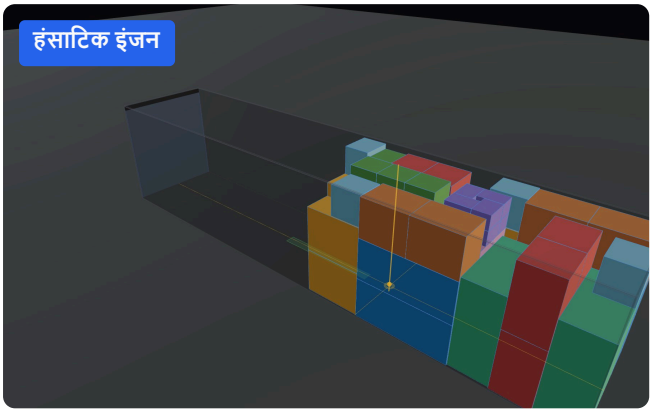
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**45 / 45**	**9.60 m (31.5 ft)**	
GPT-5.5	कानूनी	45 / 45	9.60 m (31.5 ft)	
Opus 5.5	कानूनी	45 / 45	9.61 m (31.5 ft)	
Grok 4.7	अवैध	45 / 45	10.80 m (35.4 ft)	1 गुरुत्वाकर्षण केंद्र (CoG) गलत
Kimi K3.1 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
DeepSeek V4 Pro	कानूनी	45 / 45	12.00 m (39.4 ft)	
Gemini 4 Argon	अवैध	32 / 45	7.60 m (24.9 ft)	13 गलत आइटम आकार, 1 आइटम ओवरलैप हो रहे हैं, 16 तैरते हुए आइटम
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	अवैध	10 / 45	12.00 m (39.4 ft)	1 गुरुत्वाकर्षण केंद्र (CoG) गलत

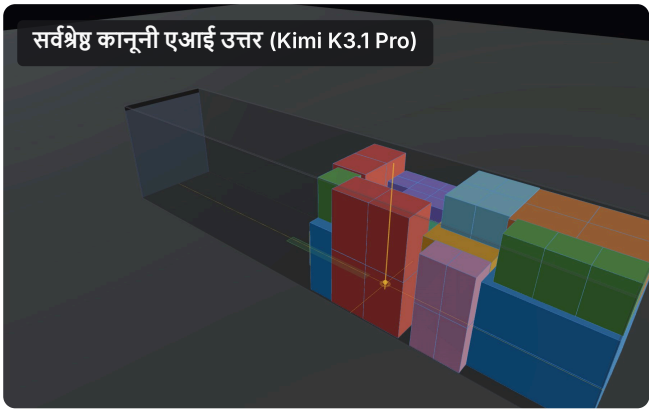
H08 सबसे छोटा लोड

वाहन	टुकड़े	उद्देश्य	परिणाम
40' High Cube	52	सबसे छोटा लोड	इंजन स्पष्ट रूप से आगे

जो इसे कठिन बनाता है

52 आइटम जो सीधे खड़े रहने चाहिए





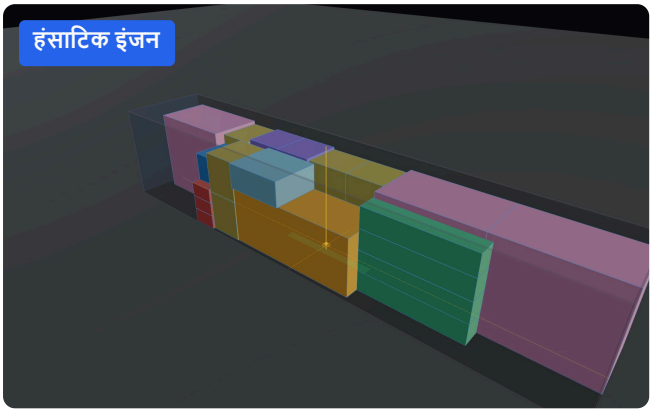
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**52 / 52**	**5.92 m (19.4 ft)**	
GPT-5.5	**कानूनी**	52 / 52	6.26 m (20.5 ft)	
Opus 5.5	**कानूनी**	52 / 52	6.20 m (20.3 ft)	
Grok 4.7	**कानूनी**	52 / 52	6.18 m (20.3 ft)	
Kimi K3.1 Pro	**कानूनी**	52 / 52	6.12 m (20.1 ft)	
DeepSeek V4 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Gemini 4 Argon	**अवैध**	51 / 52	6.51 m (21.4 ft)	1 गलत आइटम आकार, 4 आइटम ओवरलैप हो रहे हैं, 1 तैरते हुए आइटम
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	**अवैध**	52 / 52	11.25 m (36.9 ft)	14 आइटम ओवरलैप हो रहे हैं, 20 वाहन के बाहर

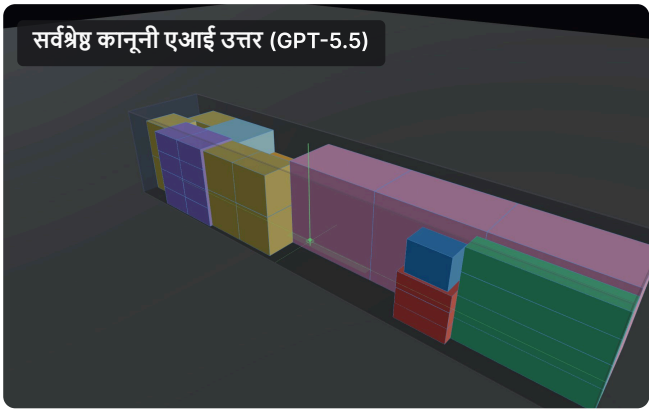
H09 सबसे छोटा लोड

वाहन	टुकड़े	उद्देश्य	परिणाम
13.6m Curtainsider	30	सबसे छोटा लोड	इंजन मामूली रूप से आगे

जो इसे कठिन बनाता है

30 आइटम जो सीधे खड़े रहने चाहिए; सबसे भारी आइटम 2,400 kg (5,291 lb)





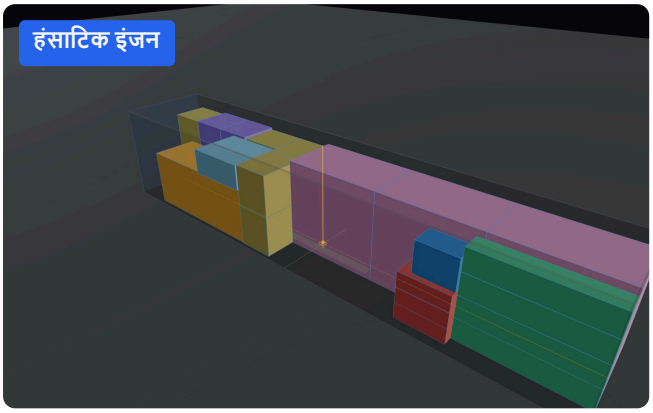
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**29 / 30**	**13.36 m (43.8 ft)**	
GPT-5.5	**कानूनी**	29 / 30	13.56 m (44.5 ft)	
Opus 5.5	**कानूनी**	29 / 30	13.60 m (44.6 ft)	
Grok 4.7	**कानूनी**	29 / 30	13.60 m (44.6 ft)	
Kimi K3.1 Pro	**कानूनी**	29 / 30	13.60 m (44.6 ft)	
DeepSeek V4 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Gemini 4 Argon	**अवैध**	19 / 30	11.60 m (38.1 ft)	11 गलत आइटम आकार, 7 आइटम ओवरलैप हो रहे हैं
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	**अवैध**	30 / 30	13.36 m (43.8 ft)	17 आइटम ओवरलैप हो रहे हैं, 25 वाहन के बाहर, 7 तैरते हुए आइटम, 1 एक्सल पर अधिक भार

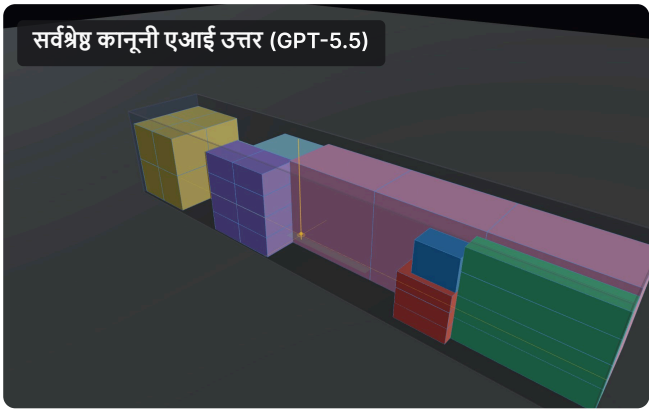
H09 सबसे कम पैलेट स्थान

वाहन	टुकड़े	उद्देश्य	परिणाम
13.6m Curtainsider	30	सबसे कम पैलेट स्थान	इंजन मामूली रूप से आगे

जो इसे कठिन बनाता है

30 आइटम जो सीधे खड़े रहने चाहिए; सबसे भारी आइटम 2,400 kg (5,291 lb)





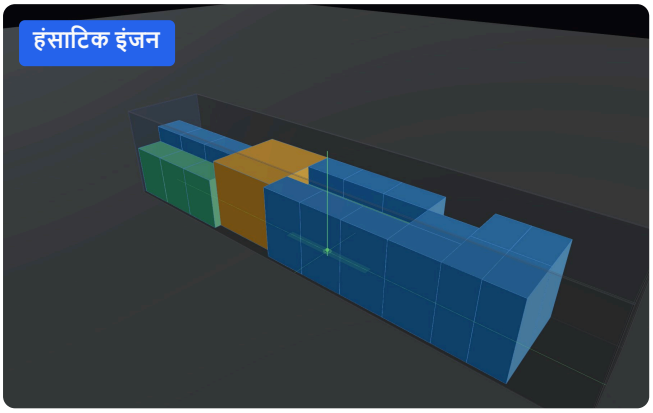
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**29 / 30**	**29 पैलेट स्थान, 13.36 m (43.8 ft)**	
GPT-5.5	**कानूनी**	29 / 30	29 पैलेट स्थान, 13.60 m (44.6 ft)	
Opus 5.5	**कानूनी**	29 / 30	31 पैलेट स्थान, 13.60 m (44.6 ft)	
Grok 4.7	**कानूनी**	29 / 30	29 पैलेट स्थान, 13.62 m (44.7 ft)	
Kimi K3.1 Pro	**कानूनी**	29 / 30	29 पैलेट स्थान, 13.60 m (44.6 ft)	
DeepSeek V4 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Gemini 4 Argon	**अवैध**	30 / 30	29 पैलेट स्थान, 11.72 m (38.5 ft)	18 आइटम ओवरलैप हो रहे हैं, 1 वाहन के बाहर, 3 दबे हुए आइटम, 1 एक्सल पर अधिक भार
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	**अवैध**	30 / 30	34 पैलेट स्थान, 13.36 m (43.8 ft)	4 आइटम ओवरलैप हो रहे हैं, 25 वाहन के बाहर, 1 एक्सल पर अधिक भार, 1 पंक्तियाँ गलत संरेखित हैं

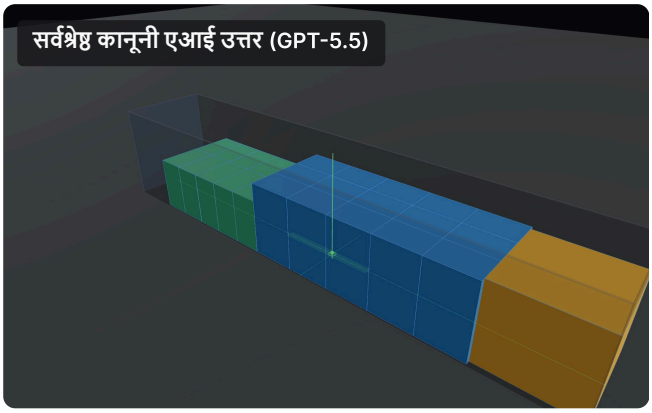
H10 सबसे छोटा लोड

वाहन	टुकड़े	उद्देश्य	परिणाम
13.6m Curtainsider	26	सबसे छोटा लोड	बराबरी

जो इसे कठिन बनाता है

26 गैर-ढेर करने योग्य आइटम; 16 आइटम जो सीधे खड़े रहने चाहिए; सबसे भारी आइटम 666 kg (1,468 lb)





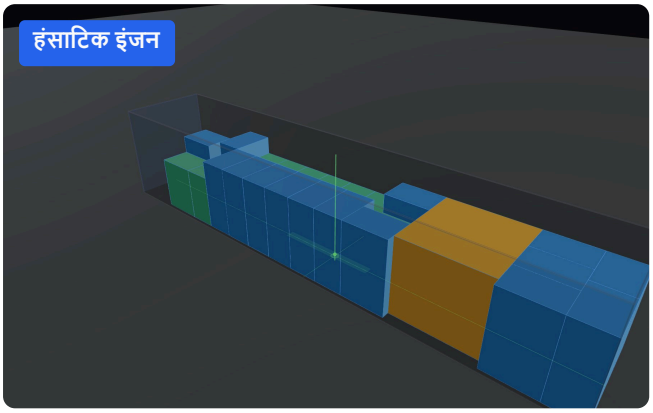
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**26 / 26**	**12.20 m (40.0 ft)**	
GPT-5.5	कानूनी	26 / 26	12.20 m (40.0 ft)	
Opus 5.5	कानूनी	26 / 26	12.20 m (40.0 ft)	
Grok 4.7	कानूनी	26 / 26	12.20 m (40.0 ft)	
Kimi K3.1 Pro	कानूनी	26 / 26	12.20 m (40.0 ft)	
DeepSeek V4 Pro	कानूनी	26 / 26	12.20 m (40.0 ft)	
Gemini 4 Argon	कानूनी	26 / 26	12.60 m (41.3 ft)	
Qwen3p8	कानूनी	26 / 26	12.20 m (40.0 ft)	
Magistral Medium	कानूनी	26 / 26	13.00 m (42.7 ft)	

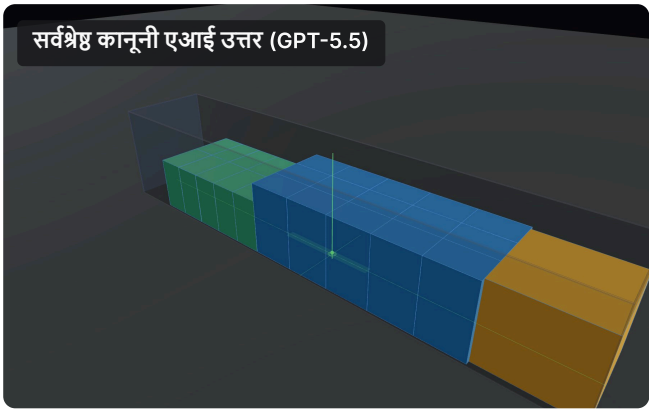
H10 सबसे कम पैलेट स्थान

वाहन	टुकड़े	उद्देश्य	परिणाम
13.6m Curtainsider	26	सबसे कम पैलेट स्थान	बराबरी

जो इसे कठिन बनाता है

26 गैर-ढेर करने योग्य आइटम; 16 आइटम जो सीधे खड़े रहने चाहिए; सबसे भारी आइटम 666 kg (1,468 lb)





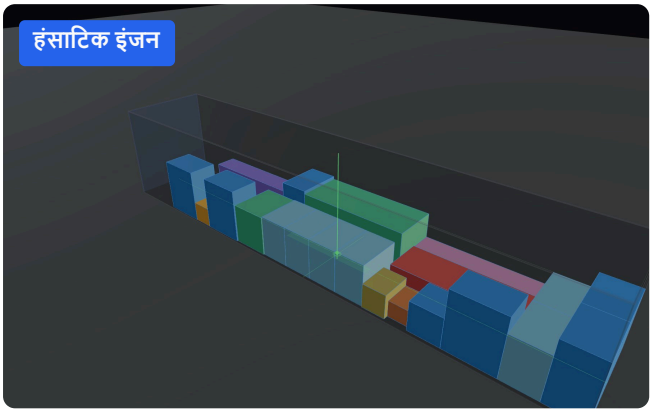
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**26 / 26**	**31 पैलेट स्थान, 12.20 m (40.0 ft)**	
GPT-5.5	कानूनी	26 / 26	31 पैलेट स्थान, 12.20 m (40.0 ft)	
Opus 5.5	कानूनी	26 / 26	31 पैलेट स्थान, 13.00 m (42.7 ft)	
Grok 4.7	कानूनी	26 / 26	31 पैलेट स्थान, 12.20 m (40.0 ft)	
Kimi K3.1 Pro	कानूनी	26 / 26	31 पैलेट स्थान, 12.20 m (40.0 ft)	
DeepSeek V4 Pro	कानूनी	26 / 26	31 पैलेट स्थान, 12.20 m (40.0 ft)	
Gemini 4 Argon	अवैध	26 / 26	31 पैलेट स्थान, 12.60 m (41.3 ft)	1 एक्सल पर अधिक भार
Qwen3p8	कानूनी	26 / 26	31 पैलेट स्थान, 13.00 m (42.7 ft)	
Magistral Medium	अवैध	26 / 26	28 पैलेट स्थान, 13.00 m (42.7 ft)	10 वाहन के बाहर

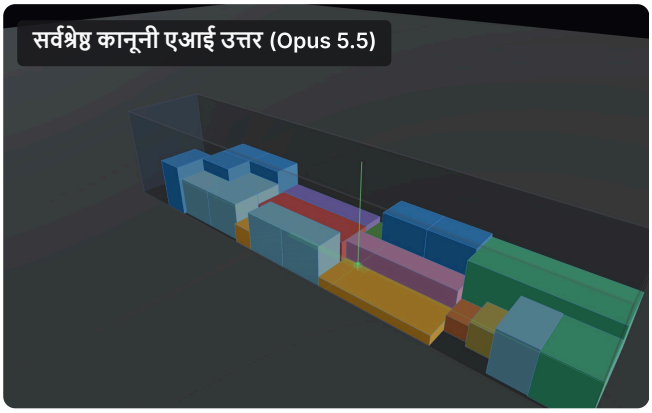
H11 सबसे छोटा लोड

वाहन	टुकड़े	उद्देश्य	परिणाम
13.6m Curtainsider	23	सबसे छोटा लोड	इंजन मामूली रूप से आगे

जो इसे कठिन बनाता है

23 गैर-ढेर करने योग्य आइटम; 23 आइटम जो सीधे खड़े रहने चाहिए; सबसे भारी आइटम 2,700 kg (5,952 lb)





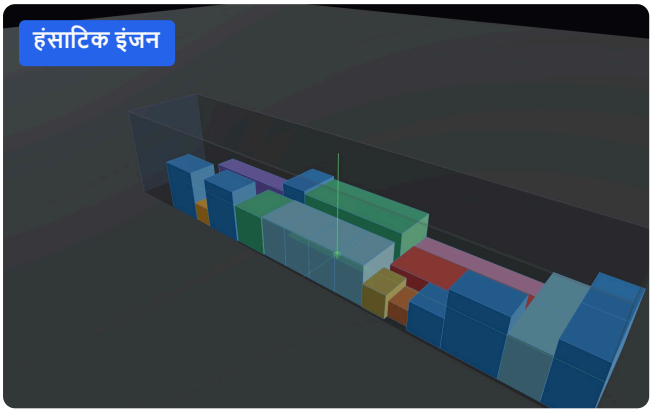
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**23 / 23**	**12.00 m (39.4 ft)**	
GPT-5.5	**कानूनी**	23 / 23	12.42 m (40.7 ft)	
Opus 5.5	**कानूनी**	23 / 23	12.20 m (40.0 ft)	
Grok 4.7	**कानूनी**	23 / 23	12.62 m (41.4 ft)	
Kimi K3.1 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
DeepSeek V4 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Gemini 4 Argon	**कानूनी**	23 / 23	13.02 m (42.7 ft)	
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	**अवैध**	23 / 23	12.34 m (40.5 ft)	7 आइटम ओवरलैप हो रहे हैं, 8 वाहन के बाहर

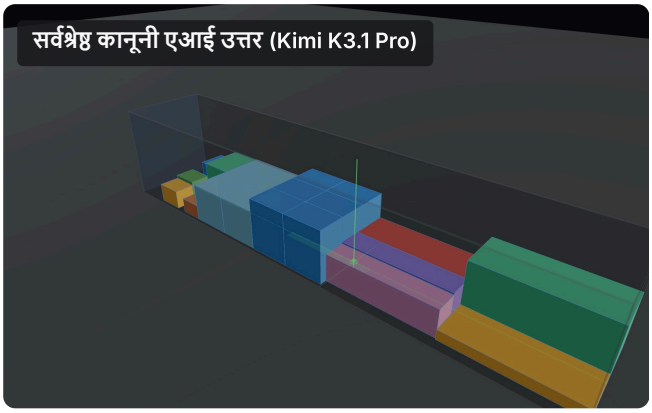
H11 सबसे कम पैलेट स्थान

वाहन	टुकड़े	उद्देश्य	परिणाम
13.6m Curtainsider	23	सबसे कम पैलेट स्थान	इंजन स्पष्ट रूप से आगे

जो इसे कठिन बनाता है

23 गैर-ढेर करने योग्य आइटम; 23 आइटम जो सीधे खड़े रहने चाहिए; सबसे भारी आइटम 2,700 kg (5,952 lb)





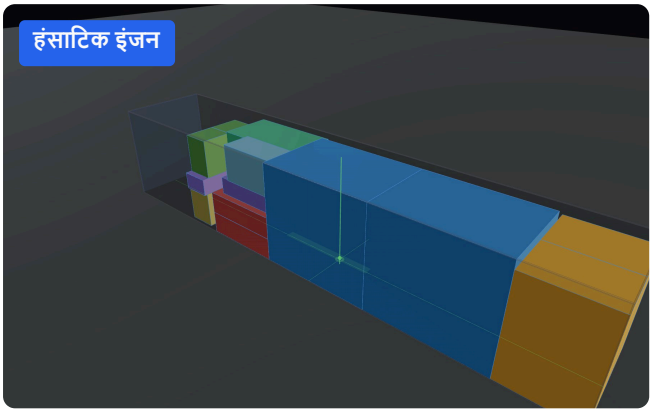
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**23 / 23**	**30 पैलेट स्थान, 12.00 m (39.4 ft)**	
GPT-5.5	**कानूनी**	23 / 23	32 पैलेट स्थान, 12.64 m (41.5 ft)	
Opus 5.5	**कानूनी**	23 / 23	32 पैलेट स्थान, 12.84 m (42.1 ft)	
Grok 4.7	प्रदाता त्रुटि	-	-	प्रदाता ने अनुरोध अस्वीकार कर दिया।
Kimi K3.1 Pro	**कानूनी**	23 / 23	32 पैलेट स्थान, 12.62 m (41.4 ft)	
DeepSeek V4 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Gemini 4 Argon	**अवैध**	22 / 23	30 पैलेट स्थान, 12.44 m (40.8 ft)	1 गलत आइटम आकार
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	**अवैध**	23 / 23	30 पैलेट स्थान, 12.60 m (41.3 ft)	11 आइटम ओवरलैप हो रहे हैं, 9 वाहन के बाहर

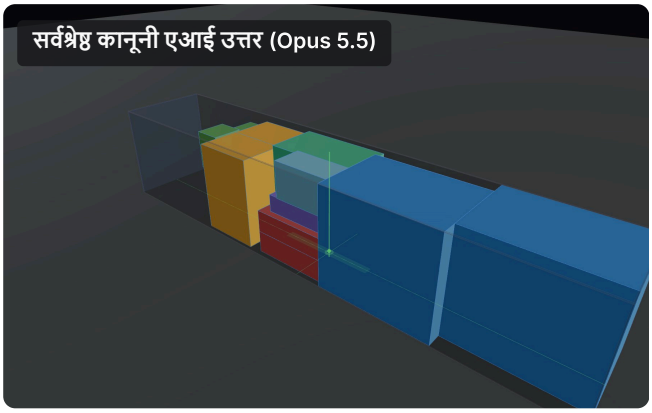
H12 सबसे छोटा लोड

वाहन	टुकड़े	उद्देश्य	परिणाम
13.6m Curtainsider	14	सबसे छोटा लोड	बराबरी

जो इसे कठिन बनाता है

6 गैर-ढेर करने योग्य आइटम; 9 आइटम जो सीधे खड़े रहने चाहिए; सबसे भारी आइटम 732 kg (1,614 lb)





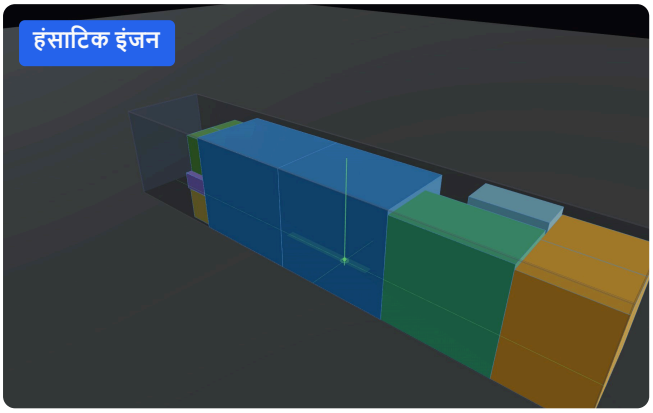
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**13 / 14**	**11.00 m (36.1 ft)**	
GPT-5.5	**कानूनी**	13 / 14	11.40 m (37.4 ft)	
Opus 5.5	**कानूनी**	13 / 14	11.00 m (36.1 ft)	
Grok 4.7	प्रदाता त्रुटि	-	-	प्रदाता ने अनुरोध अस्वीकार कर दिया।
Kimi K3.1 Pro	**कानूनी**	13 / 14	11.00 m (36.1 ft)	
DeepSeek V4 Pro	**कानूनी**	13 / 14	11.70 m (38.4 ft)	
Gemini 4 Argon	**कानूनी**	13 / 14	11.00 m (36.1 ft)	
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	**अवैध**	14 / 14	13.50 m (44.3 ft)	5 आइटम ओवरलैप हो रहे हैं, 5 वाहन के बाहर, 2 तैरते हुए आइटम

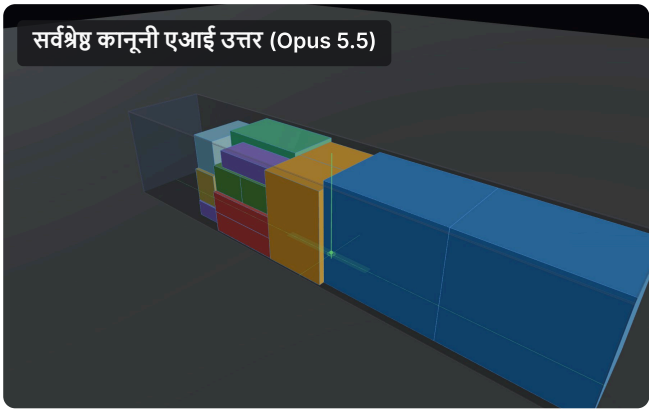
H12 सबसे कम पैलेट स्थान

वाहन	टुकड़े	उद्देश्य	परिणाम
13.6m Curtainsider	14	सबसे कम पैलेट स्थान	इंजन स्पष्ट रूप से आगे

जो इसे कठिन बनाता है

6 गैर-ढेर करने योग्य आइटम; 9 आइटम जो सीधे खड़े रहने चाहिए; सबसे भारी आइटम 732 kg (1,614 lb)





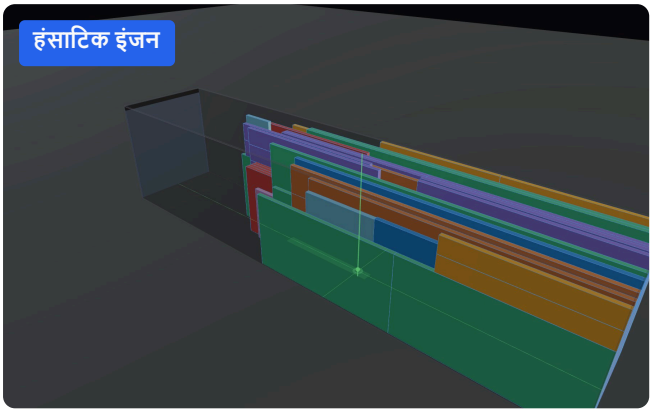
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**13 / 14**	**25 पैलेट स्थान, 11.00 m (36.1 ft)**	
GPT-5.5	**कानूनी**	13 / 14	29 पैलेट स्थान, 11.73 m (38.5 ft)	
Opus 5.5	**कानूनी**	13 / 14	28 पैलेट स्थान, 11.00 m (36.1 ft)	
Grok 4.7	प्रदाता त्रुटि	-	-	प्रदाता ने अनुरोध अस्वीकार कर दिया।
Kimi K3.1 Pro	**कानूनी**	13 / 14	28 पैलेट स्थान, 11.00 m (36.1 ft)	
DeepSeek V4 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Gemini 4 Argon	**कानूनी**	13 / 14	28 पैलेट स्थान, 11.15 m (36.6 ft)	
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	**अवैध**	14 / 14	33 पैलेट स्थान, 13.20 m (43.3 ft)	9 आइटम ओवरलैप हो रहे हैं, 9 वाहन के बाहर

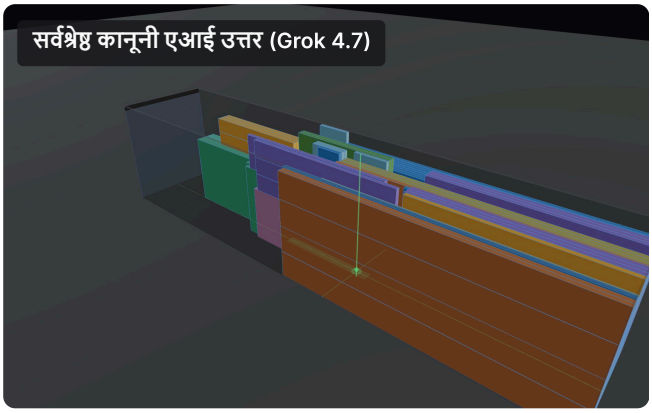
BV1 सबसे छोटा लोड

वाहन	टुकड़े	उद्देश्य	परिणाम
40' High Cube	103	सबसे छोटा लोड	इंजन स्पष्ट रूप से आगे

जो इसे कठिन बनाता है

37 आइटम जो सीधे खड़े रहने चाहिए; 8.27 m (27.1 ft) तक लंबे आइटम





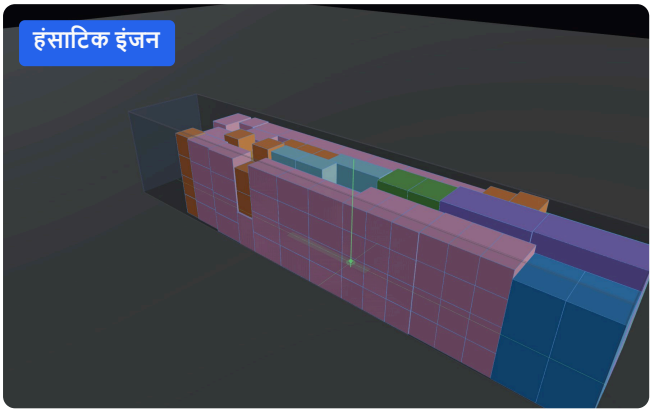
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**103 / 103**	**9.17 m (30.1 ft)**	
GPT-5.5	**कानूनी**	92 / 103	8.27 m (27.1 ft)	
Opus 5.5	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Grok 4.7	**कानूनी**	103 / 103	10.66 m (35.0 ft)	
Kimi K3.1 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
DeepSeek V4 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Gemini 4 Argon	**अवैध**	77 / 103	8.36 m (27.4 ft)	1 गलत आइटम आकार, 4 ऊंचाई सीमा से अधिक, 4 तैरते हुए आइटम
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	**अवैध**	22 / 103	16.54 m (54.3 ft)	9 वाहन के बाहर

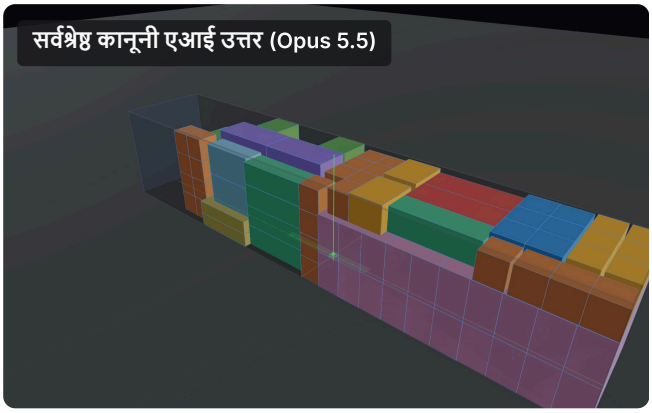
BV2 सबसे छोटा लोड

वाहन	टुकड़े	उद्देश्य	परिणाम
13.6m Curtainsider	153	सबसे छोटा लोड	इंजन स्पष्ट रूप से आगे

जो इसे कठिन बनाता है

2 गैर-ढेर करने योग्य आइटम; सबसे भारी आइटम 578 kg (1,274 lb)





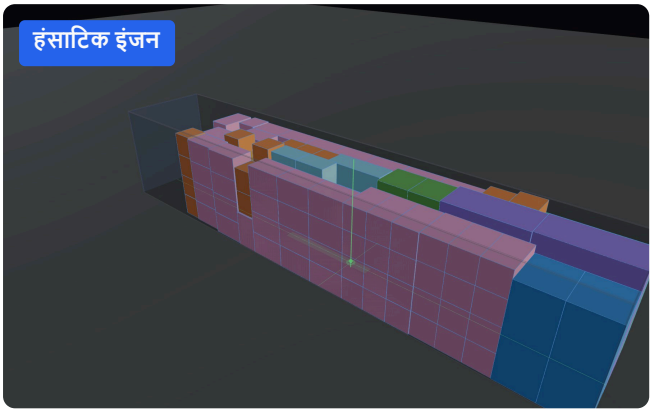
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**153 / 153**	**11.52 m (37.8 ft)**	
GPT-5.5	कानूनी	153 / 153	13.14 m (43.1 ft)	
Opus 5.5	कानूनी	153 / 153	11.78 m (38.6 ft)	
Grok 4.7	कानूनी	153 / 153	12.30 m (40.4 ft)	
Kimi K3.1 Pro	कानूनी	153 / 153	11.99 m (39.3 ft)	
DeepSeek V4 Pro	कानूनी	153 / 153	12.79 m (42.0 ft)	
Gemini 4 Argon	अवैध	147 / 153	12.77 m (41.9 ft)	6 गलत आइटम आकार
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	अवैध	56 / 153	14.02 m (46.0 ft)	39 आइटम ओवरलैप हो रहे हैं, 27 वाहन के बाहर

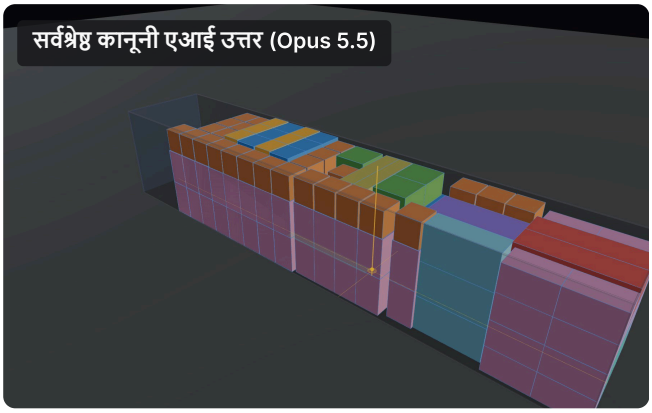
BV2 सबसे कम पैलेट स्थान

वाहन	टुकड़े	उद्देश्य	परिणाम
13.6m Curtainsider	153	सबसे कम पैलेट स्थान	इंजन मामूली रूप से आगे

जो इसे कठिन बनाता है

2 गैर-ढेर करने योग्य आइटम; सबसे भारी आइटम 578 kg (1,274 lb)





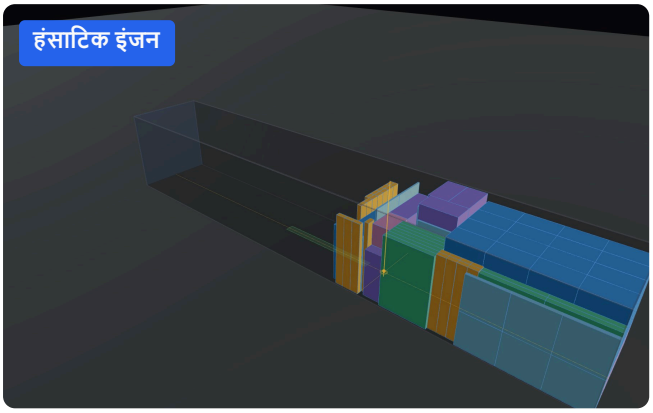
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**153 / 153**	**29 पैलेट स्थान, 11.52 m (37.8 ft)**	
GPT-5.5	**कानूनी**	153 / 153	32 पैलेट स्थान, 12.60 m (41.3 ft)	
Opus 5.5	**कानूनी**	153 / 153	29 पैलेट स्थान, 11.67 m (38.3 ft)	
Grok 4.7	**कानूनी**	153 / 153	32 पैलेट स्थान, 12.78 m (41.9 ft)	
Kimi K3.1 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
DeepSeek V4 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Gemini 4 Argon	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	**अवैध**	153 / 153	34 पैलेट स्थान, 14.82 m (48.6 ft)	7 आइटम ओवरलैप हो रहे हैं, 31 वाहन के बाहर, 86 ऊंचाई सीमा से अधिक, 96 तैरते हुए आइटम

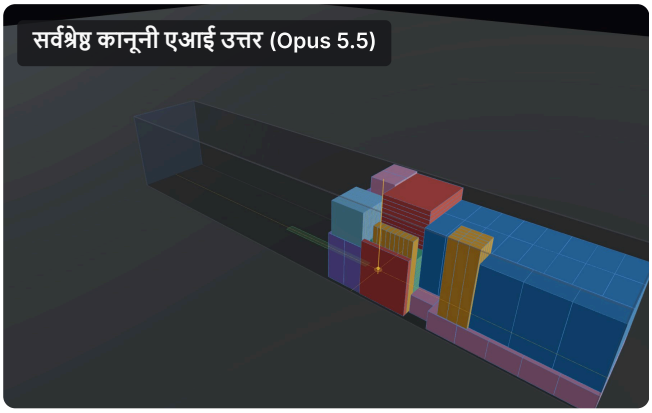
BV3 सबसे छोटा लोड

वाहन	टुकड़े	उद्देश्य	परिणाम
53ft Dry Van	119	सबसे छोटा लोड	इंजन स्पष्ट रूप से आगे

जो इसे कठिन बनाता है

19 गैर-ढेर करने योग्य आइटम; 35 आइटम जो सीधे खड़े रहने चाहिए





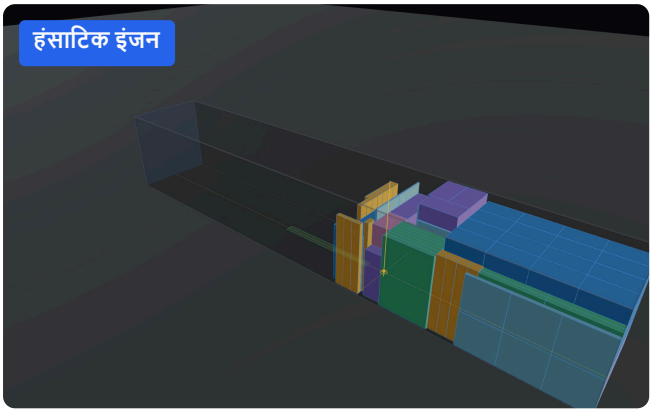
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**119 / 119**	**6.83 m (22.4 ft)**	
GPT-5.5	**कानूनी**	119 / 119	9.14 m (30.0 ft)	
Opus 5.5	**कानूनी**	119 / 119	6.99 m (22.9 ft)	
Grok 4.7	**कानूनी**	119 / 119	8.10 m (26.6 ft)	
Kimi K3.1 Pro	कोई योजना नहीं दी गई	-	-	उत्तर में कोई वैध योजना नहीं थी।
DeepSeek V4 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Gemini 4 Argon	**अवैध**	115 / 119	8.36 m (27.4 ft)	4 गलत आइटम आकार, 5 आइटम ओवरलैप हो रहे हैं, 3 दबे हुए आइटम
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	**अवैध**	22 / 119	16.76 m (55.0 ft)	11 वाहन के बाहर

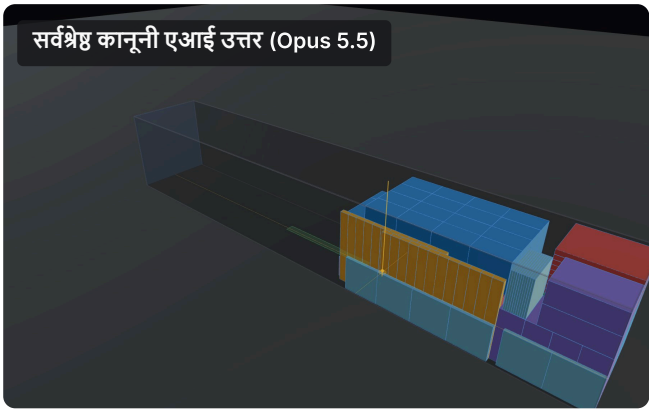
BV3 सबसे कम पैलेट स्थान

वाहन	टुकड़े	उद्देश्य	परिणाम
53ft Dry Van	119	सबसे कम पैलेट स्थान	इंजन स्पष्ट रूप से आगे

जो इसे कठिन बनाता है

19 गैर-ढेर करने योग्य आइटम; 35 आइटम जो सीधे खड़े रहने चाहिए





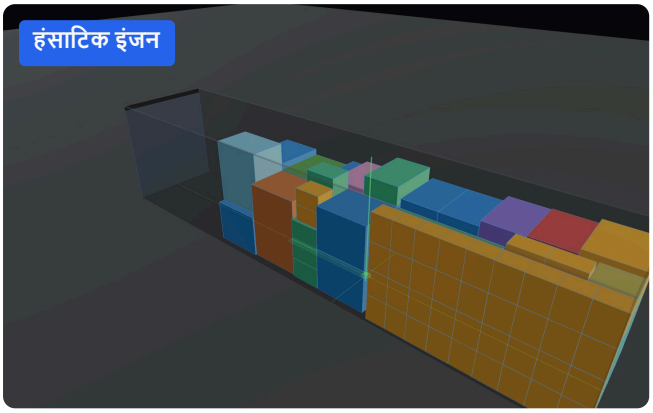
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**119 / 119**	**17 पैलेट स्थान, 6.83 m (22.4 ft)**	
GPT-5.5	**कानूनी**	119 / 119	23 पैलेट स्थान, 9.07 m (29.8 ft)	
Opus 5.5	**कानूनी**	119 / 119	18 पैलेट स्थान, 6.99 m (22.9 ft)	
Grok 4.7	**कानूनी**	119 / 119	22 पैलेट स्थान, 8.64 m (28.3 ft)	
Kimi K3.1 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
DeepSeek V4 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Gemini 4 Argon	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	**अवैध**	107 / 119	21 पैलेट स्थान, 16.18 m (53.1 ft)	12 गलत आइटम आकार, 66 आइटम ओवरलैप हो रहे हैं, 35 वाहन के बाहर, 60 तैरते हुए आइटम, 1 दबे हुए आइटम, 1 ढीले समूह जो लोड द्वारा स्थिर नहीं हैं

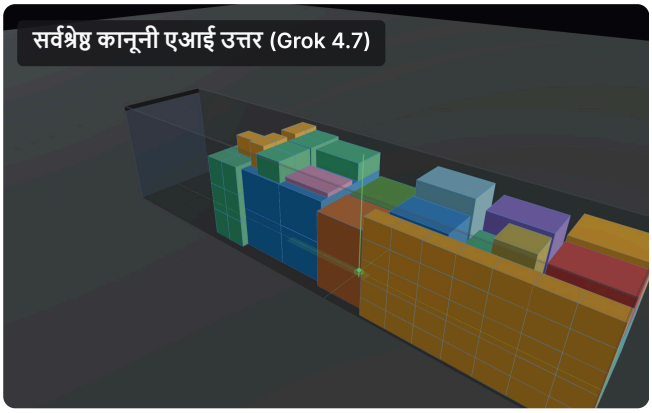
BV4 सबसे छोटा लोड

वाहन	टुकड़े	उद्देश्य	परिणाम
40' High Cube	145	सबसे छोटा लोड	इंजन स्पष्ट रूप से आगे

जो इसे कठिन बनाता है

11 गैर-ढेर करने योग्य आइटम; 11 नाजुक आइटम, केवल ऊपर; 20 आइटम जो सीधे खड़े रहने चाहिए





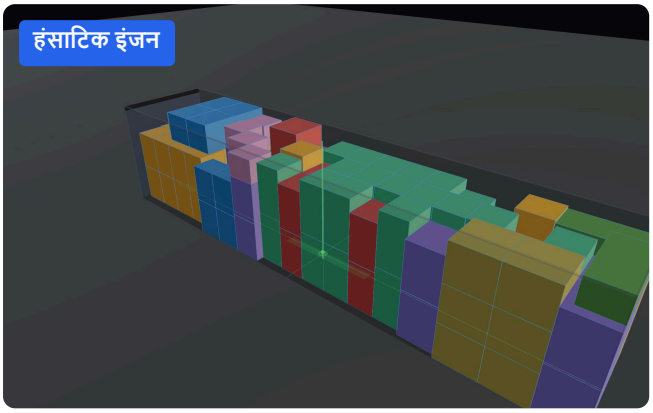
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**145 / 145**	**8.57 m (28.1 ft)**	
GPT-5.5	कानूनी	145 / 145	11.96 m (39.2 ft)	
Opus 5.5	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Grok 4.7	कानूनी	145 / 145	8.85 m (29.0 ft)	
Kimi K3.1 Pro	कानूनी	145 / 145	10.87 m (35.7 ft)	
DeepSeek V4 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Gemini 4 Argon	अवैध	145 / 145	10.05 m (33.0 ft)	19 आइटम ओवरलैप हो रहे हैं, 3 वाहन के बाहर, 9 ऊंचाई सीमा से अधिक, 4 तैरते हुए आइटम
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	कानूनी	11 / 145	10.33 m (33.9 ft)	

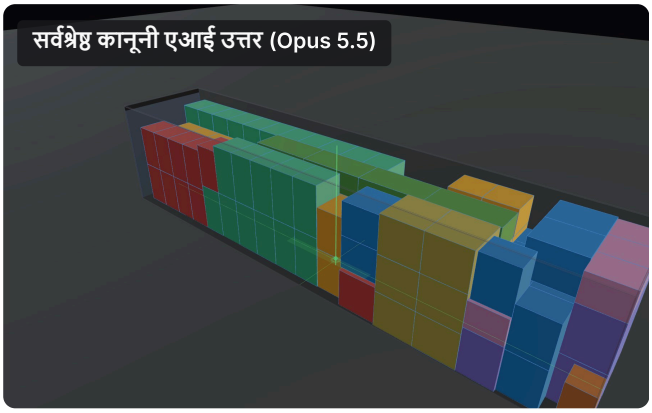
BV5 सबसे छोटा लोड

वाहन	टुकड़े	उद्देश्य	परिणाम
40' High Cube	136	सबसे छोटा लोड	इंजन स्पष्ट रूप से आगे

जो इसे कठिन बनाता है

136 आइटम जो सीधे खड़े रहने चाहिए





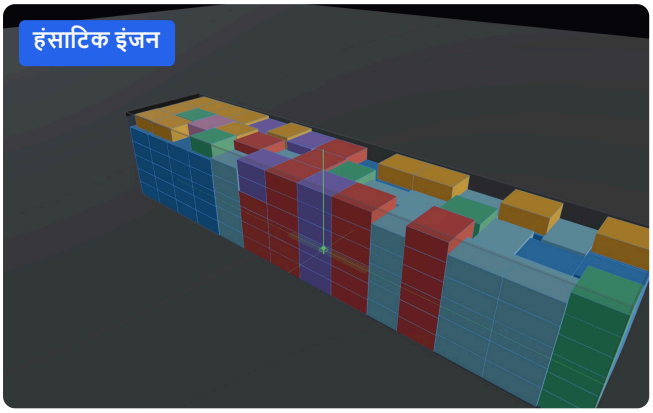
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**127 / 136**	**12.00 m (39.4 ft)**	
GPT-5.5	**कानूनी**	115 / 136	11.73 m (38.5 ft)	
Opus 5.5	**कानूनी**	124 / 136	11.77 m (38.6 ft)	
Grok 4.7	**कानूनी**	115 / 136	11.87 m (39.0 ft)	
Kimi K3.1 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
DeepSeek V4 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Gemini 4 Argon	**अवैध**	121 / 136	12.00 m (39.4 ft)	1 ऐसे आइटम को झुकाया गया जो सीधा रहना चाहिए, 12 आइटम ओवरलैप हो रहे हैं, 6 वाहन के बाहर, 6 तैरते हुए आइटम
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	**अवैध**	51 / 136	12.89 m (42.3 ft)	34 वाहन के बाहर, 6 तैरते हुए आइटम

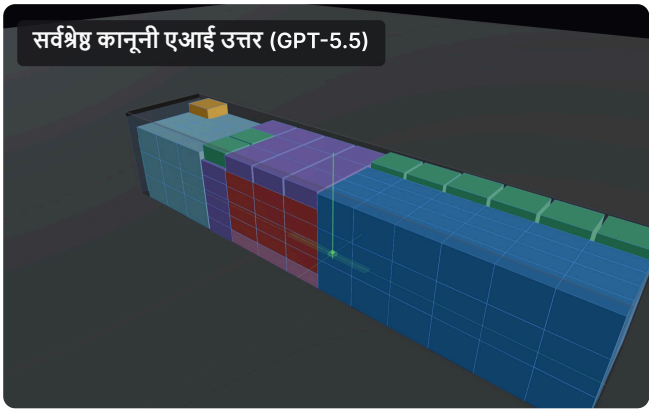
BV6 सबसे छोटा लोड

वाहन	टुकड़े	उद्देश्य	परिणाम
40' Dry Container	253	सबसे छोटा लोड	इंजन स्पष्ट रूप से आगे

जो इसे कठिन बनाता है

मिश्रित आकार, कोई विशेष संकेत नहीं



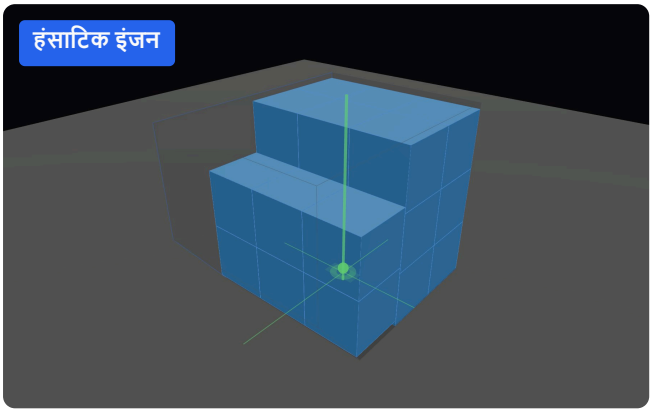


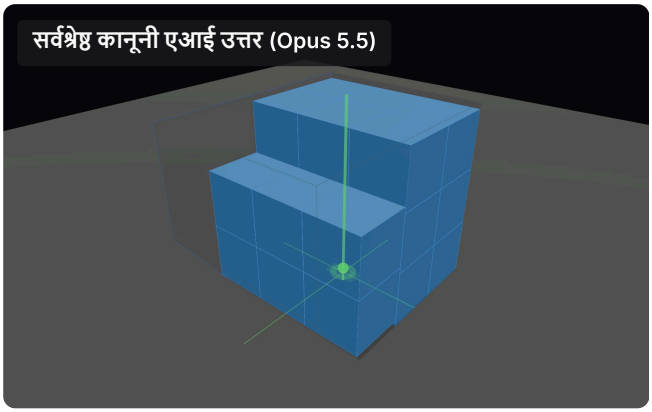
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**244 / 253**	**11.98 m (39.3 ft)**	
GPT-5.5	**कानूनी**	238 / 253	11.97 m (39.3 ft)	
Opus 5.5	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Grok 4.7	**कानूनी**	218 / 253	11.98 m (39.3 ft)	
Kimi K3.1 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
DeepSeek V4 Pro	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Gemini 4 Argon	**अवैध**	197 / 253	11.50 m (37.7 ft)	13 आइटम ओवरलैप हो रहे हैं
Qwen3p8	कोई योजना नहीं दी गई	-	-	मॉडल ने योजना लिखने से पहले अपने आउटपुट सीमा को पार कर लिया।
Magistral Medium	**अवैध**	42 / 253	12.18 m (40.0 ft)	3 वाहन के बाहर

U01 सबसे अधिक वस्तुएँ रखी गईं

वाहन	टुकड़े	उद्देश्य	परिणाम
AKE (LD3)	50	सबसे अधिक वस्तुएँ रखी गईं	बराबरी

जो इसे कठिन बनाता है
ढलान वाली दीवार





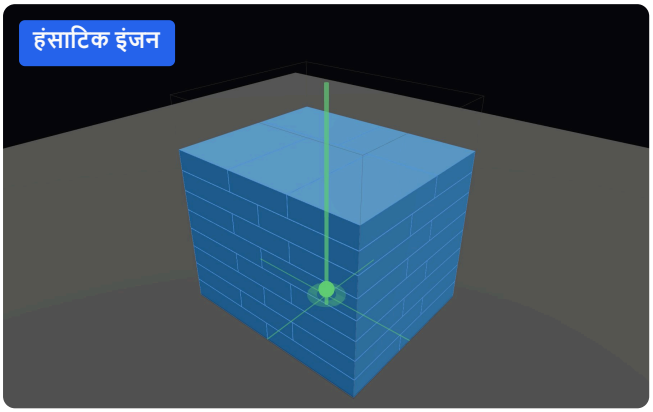
प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**24 / 50**	**24**	
GPT-5.5	कानूनी	18 / 50	18	
Opus 5.5	कानूनी	24 / 50	24	
Grok 4.7	कानूनी	24 / 50	24	
Kimi K3.1 Pro	कानूनी	24 / 50	24	
DeepSeek V4 Pro	कानूनी	18 / 50	18	
Gemini 4 Argon	कानूनी	24 / 50	24	
Qwen3p8	कानूनी	24 / 50	24	
Magistral Medium	अवैध	24 / 50	24	3 एयर कंटेनर कंटूर के बाहर

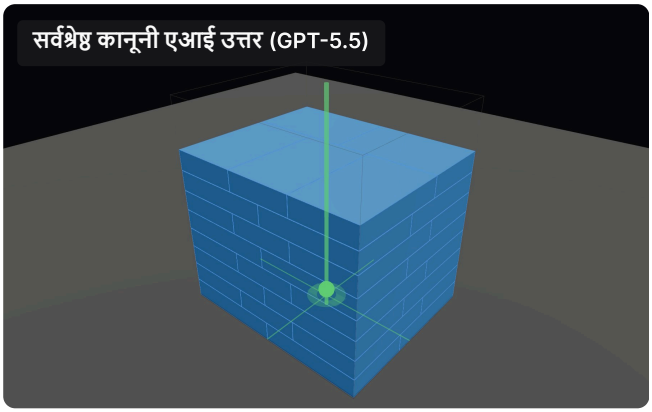
P01 सबसे अधिक इंटरलॉक्ड पैलेट

वाहन	टुकड़े	उद्देश्य	परिणाम
Industrial / Block 1200×1000	40	सबसे अधिक इंटरलॉक्ड पैलेट	बराबरी

जो इसे कठिन बनाता है

40 आइटम जो सीधे खड़े रहने चाहिए; 0.60 m (2.0 ft) तक लंबे आइटम





प्रतिभागी	स्थिति	रखे गए	स्कोर	टिप्पणियाँ
Hansatic इंजन	**कानूनी**	**40 / 40**	**100%**	
GPT-5.5	कानूनी	40 / 40	100%	
Opus 5.5	कानूनी	40 / 40	100%	
Grok 4.7	कानूनी	40 / 40	100%	
Kimi K3.1 Pro	कानूनी	40 / 40	100%	
DeepSeek V4 Pro	कानूनी	40 / 40	100%	
Gemini 4 Argon	अवैध	35 / 40	100%	5 गलत आइटम आकार, 12 तैरते हुए आइटम
Qwen3p8	कानूनी	40 / 40	100%	
Magistral Medium	अवैध	40 / 40	100%	30 आइटम ओवरलैप हो रहे हैं, 9 वाहन के बाहर