

Nadel im Heuhaufen

Laya questions on Kármán sentences

famulus-logic 2026-10-13

Table of contents

1 Question catalog	1
1.1 Logical relation	1
1.2 Sentence questions	2
1.2.1 about_person	2
1.2.2 about_named_person	2
1.2.3 mentions_date	2
2 Needle	2
3 Hay	3
3.1 Hay rows labeled positive for relation	3
3.2 Hay rows labeled positive for about_person	4
3.3 Hay rows labeled positive for about_named_person	4
3.4 Hay rows labeled positive for mentions_date	5
4 Ten sentences	5
4.1 Sentence 1	5
4.2 Sentence 2	6
4.3 Sentence 3	6
4.4 Sentence 4	7
4.5 Sentence 5	7
4.6 Sentence 6	8
4.7 Sentence 7	8
4.8 Sentence 8	8
4.9 Sentence 9	9
4.10 Sentence 10	9
5 Metrics	10
5.1 Logical relation	10
5.2 Sentence questions	10
5.3 Timing	10

1 Question catalog

Laya receives a state and typed questions. It does not generate the label. Model `aac6fef/laya-multilingual-mlx`, context `max_len 1024`, question head `head_max_len 256`. Person name for this run: **Tsien**.

The first titled line in the seed-13 shuffle is an address: Professor Theodor von Karman, Director of the Daniel Guggenheim Laboratory. It names a person and does not make a claim, so the needle is the next finite sentence, about Dr. Tsien.

1.1 Logical relation

Pair call. The examples sit in the state. The question is a 3-way choice.

Judge whether sentence S logically implies expression P.

S implies P only when P must be true whenever S is true, by the form of the sentences or by the meaning of the words. A shared topic, a shared name, or a plausible extra fact does not count.

Example where S does imply P: S: The letter was written in March 1598 in Graz.
P: The letter was written in 1598.
Example where S does not imply P, although the wording is close:
S: The letter was written in March 1598 in Graz.
P: The letter was written in March 1599 in Graz. S: {sentence} P: {probe}

```
{
  "type": "choice",
  "instructions": "Which relation holds? entails means S implies P. contradicts means S and P cannot both be true. independent means neither, including wording that is close but does not follow.",
  "criteria": [ "entails", "contradicts", "independent" ] }
```

1.2 Sentence questions

One predict when the head can hold every instruction. State is the sentence alone.

1.2.1 about_person

Yes when the sentence states something about a person: what they did, wrote, believed, or were. No when a person is only a possessor, an address line, or not present.
Positive example, the answer is yes: Theodore von Kármán wrote the introduction.
Negative example, close but the answer is no: The Kármán vortex street is stable at this Reynolds number.

1.2.2 about_named_person

Yes when the sentence states something about the person named Tsien. No when Tsien is absent, or present only as a title, a street, or a different person.
Positive example, the answer is yes: Tsien wrote the introduction.
Negative example, close but the answer is no: The Tsien vortex street is stable at this Reynolds number.

1.2.3 mentions_date

Yes when a calendar day, month, or year is mentioned. No when a number is a page, a measurement, a count, or a section.
Positive example, the answer is yes: The note is dated 12 March 1932.
Negative example, close but the answer is no: The model uses 12 modes and section 1932 of the report.

2 Needle

Key 115-2_collected_works_general_correspondence_i|10|5, language en.

Dr. Tsien and Marble expect to handle the Elasticity papers , and Dr. Millsaps volunteered for those of Mechanics and Physics , while the writer with some help expects to handle most of the general topics and Hungarian translations .

Filled state sent with the implication probe:

Judge whether sentence S logically implies expression P.

S implies P only when P must be true whenever S is true, by the form of the sentences or by the meaning of the words. A shared topic, a shared name, or a plausible extra fact does not count.

Example where S does imply P: S: The letter was written in March 1598 in Graz.

P: The letter was written in 1598.

Example where S does not imply P, although the wording is close:

S: The letter was written in March 1598 in Graz.

P: The letter was written in March 1599 in Graz.

S: Dr. Tsien and Marble expect to handle the Elasticity papers , and Dr. Millsaps volunteered for those of Mechanics and Physics , while the writer with some help expects to handle most of the general topics and Hungarian translations .

P: Dr. Tsien expects to handle the Elasticity papers.

Rule for the implication: The subject 'Dr. Tsien and Marble' distributes, so Dr. Tsien expects to handle those papers.

Probe P: Dr. Tsien expects to handle the Elasticity papers.

Prediction **contradicts** (probability 0.598). Probabilities: {"entails": 0.2547, "contradicts": 0.5977, "independent": 0.1476}.

Close miss (Those papers are the ones Dr. Millsaps volunteered for. The wording stays close and the implication fails.): Dr. Tsien expects to handle the Mechanics and Physics papers.

Prediction **contradicts** (probability 0.746). Probabilities: {"entails": 0.0966, "contradicts": 0.746, "independent": 0.1574}.

Question	Gold	Prediction	Probability
about_person	yes	no	0.897
about_named_person	yes	no	0.977
mentions_date	no	no	0.885

3 Hay

1000 other sentences, each paired with the needle probe and each asked the three sentence questions. Wall time for the hay: 394.2 s. Pair calls: 7.71 items/s, p50 0.0818 s, p95 0.3739 s. Sentence calls: 3.78 items/s, p50 0.1577 s, p95 0.8300 s.

Question	Positive answers
relation	1
about_person	12
about_named_person	11
mentions_date	113

3.1 Hay rows labeled positive for relation

- (0.466) ONE OR TWO OF THEM HAVE PREVIOUS ENGAGEMENTS THAT MAY DELAY THEIR ARRIVAL IN TACOMA .

■ 3.2 Hay rows labeled positive for about_person

- (0.537) Hierdurch würden die Brustmuskeln entlastet und , da die tragende Partie für Luft-druckschwankungen relativ unempfindlich sei , werde die Stabilität immer mehr automatisch .
- (0.518) This report is a condensed compilation of model tests and the results obtained , together with certain mathematical analysis pertaining thereto .
- (1.000) PAGE THREE LINE TWELVE k EQUALS POINT ZERO ZERO THREE FOUR THREE V NAUGHT .
- (0.609) I sincerely hope that your visit in Paris has been both profitable and rest ful .
- (0.508) Man kann nur das verantworten , was man beurteilen kann , und man kann nur beurteilen , was man versteht .
- (0.510) The variables x and y are length coordinates along the center line , U is the velocity , ρ is the density of air in the undisturbed flow .
- (0.800) Of these , the one letter from Prof . Knapp complying with Section 11 of the contract requests notification to proceed ; the letter signed by Profs .
- (0.612) This means that , in the reduction of displacement , less and less accuracy is obtained as the ground movements slow up , or periods increase .
- (0.582) Die Arbeitslosigkeit wächst wieder , trotz des allmählich lächerlich klein gewordenen Löhne .
- (0.744) Wird die Wärme durch Materie über - tragen , so sprechen wir von Wärmeleitung , bzw .
- (0.552) The discussion of frequency variation developed in article “ c “ of Section 35 suggested the probable importance of a clearer understanding of the significance of this factor .
- (0.718) A csavarszárnnyú repülőgép feltalálása régóta foglalkoztatja az emberiséget .

■ 3.3 Hay rows labeled positive for about_named_person

- (0.543) A method is presented for predicting the behavior of stress waves in a bar of material whose stress - strain relation is partly or entirely concave upward .
- (0.523) Potenz des Auftriebes .Da Minimum der ersten Groesse bestimmt den flachen Gleitwinkel des Flugzeuges , das Minimum der Flugzahl ist massgebend fuer die kleinste Sinkgeschwindigkeit , also fuer Steigleistung und Segelfachigkeit .
- (0.850) But integrals of the Bessel equation of higher order have no essential singularity .
- (0.551) Space limitations in the wind tunnel required the side spans to be 16 inches shorter than was required for exact correspondence to the prototype .
- (0.602) With regard to the difference between accelerating and decelerating , he said it was necessary to define physically what one meant .
- (0.570) Ferner muss der turbulente Strömungszustand wahrscheinlich auch von dem Standpunkt der Schwankungstheorie aus , beleuchtet werden ; es müssen die Häufigkeitsverteilung und Störungen , ihre Korrelationen ermittelt werden , vielleicht unter Heranziehung der Wahrscheinlichkeitsrechnung , wie Herr Burgers und ich in mehreren Arbeiten angestrebt haben .
- (0.998) PAGE THREE LINE TWELVE k EQUALS POINT ZERO ZERO THREE FOUR THREE V NAUGHT .
- (0.535) Since both Dr. Bollay and Dr. Rannie encouraged me to think that you might be interested in the diabatic flow work , I have enclosed a copy of the paper which I should be glad to have you keep .
- (0.799) Of these , the one letter from Prof . Knapp complying with Section 11 of the contract requests notification to proceed ; the letter signed by Profs .
- (0.851) The present writer applied the recently developed law of turbulence based on the similarity principle , to the problem of skin friction and found good accordance at the highest Reynolds Numbers reached in the tests .
- (0.666) Wird die Wärme durch Materie über - tragen , so sprechen wir von Wärmeleitung , bzw .

3.4 Hay rows labeled positive for mentions_date

Showing 15 of 113.

- (0.502) NOTICE : THIS PHOTOGRAPH OR ANY PART THEREOF IS NOT TO BE REPRODUCED WITHOUT WRITTEN PERMISSION OF THE AIRESEARCH MANUFACTURING COMPANY .
- (0.628) Both of these curves are based on a maximum generator output of 1250 KW ; similar curves will be drawn based on generator output of 1000 KW and 1500 KW .
- (0.594) In conclusion it is necessary to point out that there is a considerable difference between a study of this type and one which considers some characteristic of flow in cases which are geometrically similar .
- (0.814) This is verified roughly by the first impulse on the record which should not be greatly affected by resonance) which is about the amplitude of the waves at their maximum .
- (0.609) The magnitude of this transfer of momentum is given by $-\rho \bar{u'v'}$, where ρ denotes the density of the fluid , u' the velocity fluctuation in the direction of flow , v' the velocity fluctuation perpendicular to the plane considered , and the bar indicates mean values with respect to time .
- (0.580) An extension of the theory to supercritical flow by the introduction of the equivalent eddy viscosity , as Dean suggests , has not been found fruitful , no doubt because the range of Reynolds numbers for which the molecular as well as the turbulent shear stresses are factors of major importance can not be dealt with effectively by such a simple device .
- (0.776) In this particular case (20° half vertex angle) , the shock wave is first attached at $M = 1.18$.
- (0.632) As long as s remains finite the system retains n degrees of freedom , but as $s \rightarrow \infty$, one degree “freezes” .
- (0.674) Da ich die Ferien in Budapest verbringe , bitte ich die Patentschriften unmittelbar nach dort zu senden .
- (0.534) This shift would have been objectionable , since the variable unbalance accompanying it would have been considerably greater than the limits of accuracy of the torque measurements desired .
- (0.732) Si U est la vitesse caractéristique , telle que la vitesse de le déplacement , l'échelle de les temps de le phénomène est donnée par L/U . Enfin , soient ρ le densité et μ le coefficient de viscosité de le fluide .
- (0.548) A legitim kereskedelem , amely eddig is foglalkozott paprikával és súlyos áldozatokat hozott érte , nagy érdemeket szerzett és senki sem akarja szerzett jogait sérteni .
- (0.602) A jet of cross section a is injected into the tube .
- (0.567) However , we can be guided by the fact that H_p and η^2 are very nearly proportional to p° .
- (0.526) In dem Leerlaufzug dagegen , in dem vor allem bei der Landung ein bequemes Erreichen des Höchst - auftriebes notwendig wäre , bedarf es hierzu sehr großer Ruder - ausschläge .

4 Ten sentences

Gold is constructed from the sentence. It is not a human review.

4.1 Sentence 1

116-12_reprint_requests|68|9 (en)

Please accept my thanks and sincerest regards .

Implication (Asking someone to accept one's thanks entails that the writer thanks the addressee.): The writer thanks the addressee.

Gold entails, prediction entails, probability 0.563, {"entails": 0.5634, "contradicts": 0.2638, "independent": 0.1728}.

Close miss (The direction of thanks is reversed.): The addressee thanks the writer.

Gold `independent`, prediction `contradicts`, probability 0.374, {"entails": 0.3398, "contradicts": 0.3736, "independent": 0.2866}.

Question	Gold	Prediction	Probability
<code>about_person</code>	yes	no	0.966
<code>about_named_person</code>	no	no	0.974
<code>mentions_date</code>	no	no	0.993

4.2 Sentence 2

13-32_horn_f_1924_1935|9|16 (de)

Weinblum , der übrigens jetzt auch Privatdozent an unserer Hochschule ist , gebeten , Ihnen seine sämtlichen Arbeiten über Wellenwiderstand zuzusenden , und ich nehme an , dass er dies inzwischen wohl schon getan haben wird .

Implication (The finite clause states that assumption; the relative clause about the lectureship is dropped.): Der Schreiber nimmt an, dass Weinblum die Arbeiten über Wellenwiderstand bereits zugesandt hat.

Gold `entails`, prediction `entails`, probability 0.529, {"entails": 0.5294, "contradicts": 0.4325, "independent": 0.0382}.

Close miss (Wellenwiderstand is replaced by Auftrieb. The sentence does not say the lift papers were sent.): Der Schreiber nimmt an, dass Weinblum die Arbeiten über Auftrieb bereits zugesandt hat.

Gold `independent`, prediction `contradicts`, probability 0.712, {"entails": 0.0974, "contradicts": 0.7123, "independent": 0.1902}.

Question	Gold	Prediction	Probability
<code>about_person</code>	yes	no	0.949
<code>about_named_person</code>	no	no	0.945
<code>mentions_date</code>	no	no	0.948

4.3 Sentence 3

109-7_patents_patent_and_inventions_by_others_ii|60|22 (en)

Total heat emissivity , per hour , for " Perfect " Ravine , from about 1 hour after sunrise until about 1 hour after sunset - 300 to 500 Million B.T.U.

Implication (A stated range of 300 to 500 entails the lower bound.): The heat emissivity for Perfect Ravine is at least 300 million B.T.U. per hour.

Gold `entails`, prediction `entails`, probability 0.467, {"entails": 0.4665, "contradicts": 0.4656, "independent": 0.0679}.

Close miss (The unit is B.T.U. A different unit does not follow.): The heat emissivity for Perfect Ravine is reported in kilowatts.

Gold `independent`, prediction `contradicts`, probability 0.984, {"entails": 0.0105, "contradicts": 0.9838, "independent": 0.0057}.

Question	Gold	Prediction	Probability
about_person	no	no	0.754
about_named_person	no	no	0.815
mentions_date	no	no	0.516

4.4 Sentence 4

100-5_bibliography_of_aerodynamics_topics_ca_1925_i|37|19 (de)

Castigliano 's Theorie der kleinsten Arbeit und der Grundsatz von St. Venant .

Implication (The title Theorie der kleinsten Arbeit says what the theory is about.): Castiglianos Theorie betrifft die kleinste Arbeit.

Gold entails, prediction entails, probability 0.996, {"entails": 0.9961, "contradicts": 0.0021, "independent": 0.0018}.

Close miss (Wave resistance is not the subject named in the line.): Castiglianos Theorie betrifft den Wellenwiderstand.

Gold independent, prediction contradicts, probability 0.674, {"entails": 0.0197, "contradicts": 0.6741, "independent": 0.3062}.

Question	Gold	Prediction	Probability
about_person	no	no	0.871
about_named_person	no	no	0.845
mentions_date	no	no	0.911

4.5 Sentence 5

130-28_papers_by_other_writers_schade_henry_a_i|6|23 (de)

Wir rechnen den Wasserdruck als von unten nach oben wir - kend , entsprechend der Wirklichkeit ; die Durchbiegung wird pos i- tiv nach oben angenommen ; dadurch stimmen die Vorzeichen von Durchbiegungen , Kräften , Momenten , Krümmungen usw . rechnungs - mässig mit der üblichen Trägertheorie überein .

Implication (That is the opening clause; the sign convention that follows is dropped.): Der Wasserdruck wird als von unten nach oben wirkend gerechnet.

Gold entails, prediction entails, probability 0.951, {"entails": 0.9512, "contradicts": 0.0355, "independent": 0.0133}.

Close miss (No unit is given. The added unit does not follow.): Der Wasserdruck wird in Tonnen pro Quadratmeter gerechnet.

Gold independent, prediction contradicts, probability 0.724, {"entails": 0.0797, "contradicts": 0.724, "independent": 0.1963}.

Question	Gold	Prediction	Probability
about_person	yes	no	0.898
about_named_person	no	no	0.791
mentions_date	no	yes	0.645

4.6 Sentence 6

102-10_bridges_suspension_iii|60|9 (en)

Denoting by (a) the modes involving center span only and by (b) the modes involving side spans only .

Implication (That is the definition the sentence gives for (a).): Mode (a) involves the center span only.

Gold entails, prediction contradicts, probability 0.698, {"entails": 0.1897, "contradicts": 0.6984, "independent": 0.1119}.

Close miss (Nothing is said about torsion. The label (a) is kept and a new claim is added.): Mode (a) is the first torsional mode.

Gold independent, prediction contradicts, probability 0.488, {"entails": 0.0542, "contradicts": 0.4879, "independent": 0.4579}.

Question	Gold	Prediction	Probability
about_person	no	no	0.919
about_named_person	no	no	0.925
mentions_date	no	no	0.760

4.7 Sentence 7

120-9_manuscripts_and_notes_technical_1947_i|47|7 (en)

The considerations given in the last section resulted from the authors discussions with H. L. Dryden and .

Implication (The dangling 'and' is dropped. The discussions with Dryden remain.): The considerations in the last section resulted from discussions with H. L. Dryden.

Gold entails, prediction contradicts, probability 0.754, {"entails": 0.144, "contradicts": 0.7541, "independent": 0.1019}.

Close miss (Dryden is replaced by Tsien, who is not the person named here.): The considerations in the last section resulted from discussions with Dr. Tsien.

Gold independent, prediction contradicts, probability 0.989, {"entails": 0.007, "contradicts": 0.9889, "independent": 0.0041}.

Question	Gold	Prediction	Probability
about_person	yes	no	0.812
about_named_person	no	no	0.830
mentions_date	no	no	0.901

4.8 Sentence 8

12-19_harrington_r_paul_1950_1962|32|6 (en)

I also wish to thank you again for the many kind things you have done for me in the last thirty years .

Implication (Thanks for those things entail that they were done.): The addressee has done kind things for the writer over thirty years.

Gold entails, prediction entails, probability 0.664, {"entails": 0.6638, "contradicts": 0.2086, "independent": 0.1276}.

Close miss (The direction of thanks is reversed.): The addressee thanks the writer for thirty years of kindness.

Gold `independent`, prediction `entails`, probability 0.736, {"entails": 0.7358, "contradicts": 0.2388, "independent": 0.0254}.

Question	Gold	Prediction	Probability
<code>about_person</code>	yes	no	0.890
<code>about_named_person</code>	no	no	0.977
<code>mentions_date</code>	no	no	0.980

4.9 Sentence 9

106-16_hydraulics_general_correspondence_i|15|9 (en)

It is expected that completion of testing on the Newport News model will require about ten days in February .

Implication (About ten days entails some days, and the month stays February.): Testing of the Newport News model is expected to take some days in February.

Gold `entails`, prediction `entails`, probability 0.970, {"entails": 0.9696, "contradicts": 0.0034, "independent": 0.0269}.

Close miss (February is replaced by March.): Testing of the Newport News model is expected to take about ten days in March.

Gold `independent`, prediction `contradicts`, probability 1.000, {"entails": 0.0, "contradicts": 0.9999, "independent": 0.0001}.

Question	Gold	Prediction	Probability
<code>about_person</code>	no	no	0.983
<code>about_named_person</code>	no	no	0.948
<code>mentions_date</code>	yes	no	0.790

4.10 Sentence 10

120-8_manuscripts_and_notes_technical_1947_ii|27|5 (en)

In the foregoing sections two general methods of computing flow produced by slender bodies and airfoils were presented : the method of sources and that of the Fourier integrals .

Implication (The sentence presents two methods, and the method of sources is one of them.): A method of sources was presented for computing flow.

Gold `entails`, prediction `entails`, probability 0.881, {"entails": 0.881, "contradicts": 0.0885, "independent": 0.0305}.

Close miss (Vortices are not among the methods named.): A method of vortices was presented for computing flow.

Gold `independent`, prediction `contradicts`, probability 0.992, {"entails": 0.0009, "contradicts": 0.9923, "independent": 0.0068}.

Question	Gold	Prediction	Probability
about_person	no	no	0.966
about_named_person	no	no	0.953
mentions_date	no	yes	0.581

5 Metrics

5.1 Logical relation

Twenty-two pairs: the needle, its close miss, and two probes for each of the ten sentences. Positive label entails.

n	Accuracy	Precision	Recall	F1	Trap rate
22	0.364	0.889	0.727	0.800	0.091

Trap rate is the share of close misses labeled entails.

5.2 Sentence questions

Eleven sentences (the needle and the ten). Positive label yes.

Question	n	Accuracy	Precision	Recall	F1
about_person	11	0.455	—	0.000	—
about_named_person	11	0.909	—	0.000	—
mentions_date	11	0.727	0.000	0.000	—

5.3 Timing

Pair timing covers every relation call, including the hay. Sentence timing covers every bundled sentence call.

Call	n	Wall s	Items/s	p50 s	p95 s
pair	1022	138.57	7.38	0.0795	0.3739
sentence	1011	265.61	3.81	0.1570	0.8300