

Onder de motorkap van ChatGPT

Prof. dr. Paul Van Eecke

Artificial Intelligence Laboratory - Vrije Universiteit Brussel

*Seminarie AI voor leerkrachten secundair onderwijs
Brussel - 3 december 2025*



ARTIFICIAL
INTELLIGENCE
RESEARCH GROUP



Co-funded by
the European Union

Klasse, 28/10/2024

KLASSE

 Zoek |  Login

AI in onderwijs: nuttige helper of vervangleraar?

28 oktober 2024 • 9 minuten lezen

De Tijd, 04/12/2024

Studie: 'Generatieve AI kan Belgische economie 50 miljard in tien jaar opleveren'

VRT NWS, 29/03/2024

ChatGPT doet het beter dan laatstejaarsstudenten geneeskunde



VRT NWS, 20/02/2025

Het probleem van hallucinaties fundamenteel oplossen? Dat kunnen we niet

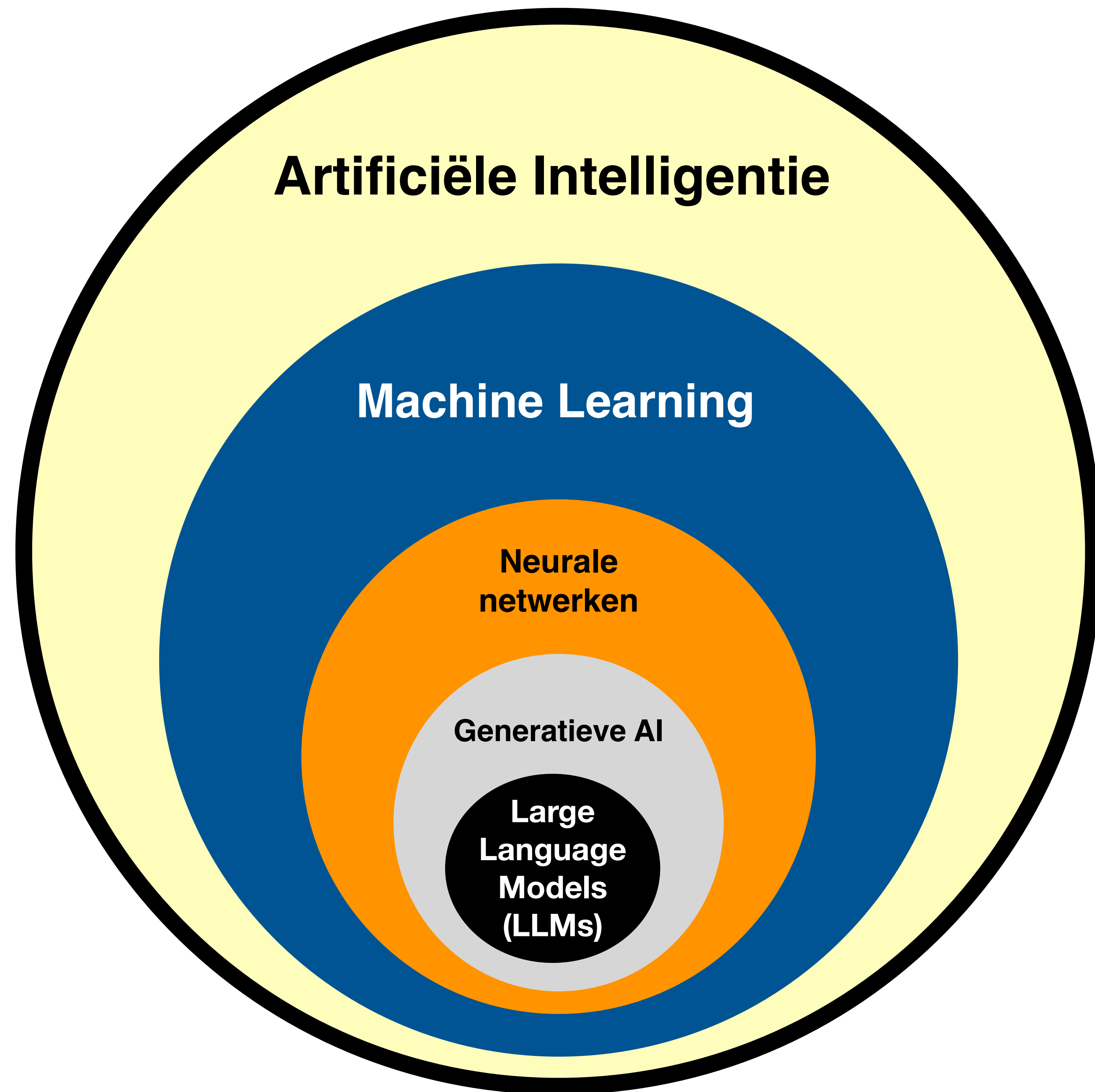
Steven Latré, hoofd Artificiële Intelligentie bij imec

De Tijd, 04/12/2024

Blind enthousiasme voor AI wordt vaak duur betaald

De Tijd, 28/12/2024

Botst AI op haar limieten in 2025?





ChatGPT


Gemini

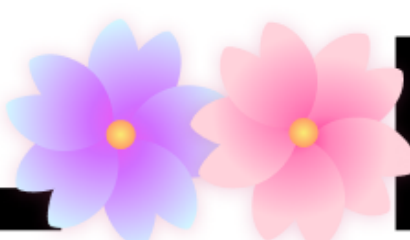


Claude

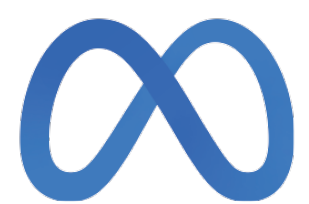


**MISTRAL
AI_**

a BigScience initiative

BL  **M**

176B params · 59 languages · Open-access

LLaMA
by  **Meta**



Grok

Large Language Models

1. Taalmodel

- Model van de waarschijnlijkheid waarmee *tokens* (~woorden) op elkaar volgen

2. Distributionele semantiek

- Voorstelling van woorden in functie van de waarschijnlijkheid waarmee andere woorden in hun buurt voorkomen

Taalmodel

Model van de waarschijnlijkheid waarmee woorden op elkaar volgen

De koekjes die je gebakken hebt zijn lekker [.....]

- knapperig
 - krokant
 - zacht
 - zoet
 - ?*blauw*
 - *~~koelkast~~
 - *~~filosofeert~~
- P(knapperig | de, koekjes, die, je, gebakken, hebt, zijn, zo, lekker)*
- ...
- P(filosofeert | de, koekjes, die, je, gebakken, hebt, zijn, zo, lekker)*

Taalmodel

Model van de waarschijnlijkheid waarmee woorden op elkaar volgen

- Meld aan je hier

$P(\text{meld, aan, je, hier})$

- Meldt je hier aan

$P(\text{meldt, je, hier, aan})$

- Meld je hier aan

$P(\text{meld, je, hier, aan})$

Taalmodel

Model van de waarschijnlijkheid waarmee woorden op elkaar volgen

Hoe bereken je de waarschijnlijkheid van een woord gegeven de voorgaande woorden?

Data!

$$P(\textit{knapperig} \mid \textit{de, koekjes, zijn, zo, lekker}) = \frac{C(\textit{de, koekjes, zijn, zo, lekker, knapperig})}{C(\textit{de, koekjes, zijn, zo, lekker})}$$

En van een reeks woorden? Kettingregel...

$$\begin{aligned} P(\textit{meld, je, hier, aan}) &= P(\textit{meld}) \times P(\textit{je} \mid \textit{meld}) \times P(\textit{hier} \mid \textit{meld, je}) \times P(\textit{aan} \mid \textit{meld, je, hier}) \\ &= \frac{C(\textit{meld})}{N} \times \frac{C(\textit{meld, je})}{C(\textit{meld})} \times \frac{C(\textit{meld, je, hier})}{C(\textit{meld, je})} \times \frac{C(\textit{meld, je, hier, aan})}{C(\textit{meld, je, hier})} \end{aligned}$$

Maar wat als de zin niet in de data voorkomt?

Taalmodel

Model van de waarschijnlijkheid waarmee woorden op elkaar volgen

Markoviaanse aanname: conditionele onafhankelijkheid

1-gram

$$P(\textit{meld}, \textit{je}, \textit{hier}, \textit{aan}) \approx P(\textit{meld}) \times P(\textit{je}) \times P(\textit{hier}) \times P(\textit{aan})$$

2-gram

$$P(\textit{meld}, \textit{je}, \textit{hier}, \textit{aan}) \approx P(\textit{meld}) \times P(\textit{je} \mid \textit{meld}) \times P(\textit{hier} \mid \textit{je}) \times P(\textit{aan} \mid \textit{hier})$$

3-gram

$$P(\textit{meld}, \textit{je}, \textit{hier}, \textit{aan}) \approx P(\textit{meld}) \times P(\textit{je} \mid \textit{meld}) \times P(\textit{hier} \mid \textit{meld}, \textit{je}) \times P(\textit{aan} \mid \textit{je}, \textit{hier})$$

Taalmodel

Model van de waarschijnlijkheid waarmee woorden op elkaar volgen

Markoviaanse aanname: conditionele onafhankelijkheid

1-gram

$$P(meld, je, hier, aan) \approx \frac{C(meld)}{N} \times \frac{C(je)}{N} \times \frac{C(hier)}{N} \times \frac{C(aan)}{N}$$

2-gram

$$P(meld, je, hier, aan) \approx \frac{C(meld)}{N} \times \frac{C(meld, je)}{C(meld)} \times \frac{C(je, hier)}{C(je)} \times \frac{C(hier, aan)}{C(hier)}$$

3-gram

$$P(meld, je, hier, aan) \approx \frac{C(meld)}{N} \times \frac{C(meld, je)}{C(meld)} \times \frac{C(meld, je, hier)}{C(meld, je)} \times \frac{C(je, hier, aan)}{C(je, hier)}$$

Taalmodel

Model van de waarschijnlijkheid waarmee woorden op elkaar volgen

	i	want	to	eat	chinese	food	lunch	spend
i	0.002	0.33	0	0.0036	0	0	0	0.00079
want	0.0022	0	0.66	0.0011	0.0065	0.0065	0.0054	0.0011
to	0.00083	0	0.0017	0.28	0.00083	0	0.0025	0.087
eat	0	0	0.0027	0	0.021	0.0027	0.056	0
chinese	0.0063	0	0	0	0	0.52	0.0063	0
food	0.014	0	0.014	0	0.00092	0.0037	0	0
lunch	0.0059	0	0	0	0	0.0029	0	0
spend	0.0036	0	0.0036	0	0	0	0	0

Taalmodel

Model van de waarschijnlijkheid waarmee woorden op elkaar volgen

De koekjes die je gebakken hebt zijn lekker [.....]

- | | |
|----------------------------|--------------|
| - knapperig | $P = 0.15$ |
| - krokant | $P = 0.13$ |
| - zacht | ... |
| - zoet | |
| - ? <i>blauw</i> | $P = 0.02$ |
| - * koelkast | ... |
| - * filosofeert | $P = 0.0001$ |

Taalmodel

Model van de waarschijnlijkheid waarmee woorden op elkaar volgen

Text genereren: woorden selecteren volgens hun waarschijnlijkheid

1
gram

Months the my and issue of year foreign new exchange's september
were recession exchange new endorsed a acquire to six executives

2
gram

Last December through the way to preserve the Hudson corporation N.
B. E. C. Taylor would seem to complete the major central planners one
point five percent of U. S. E. has already old M. X. corporation of living
on information such as more frequently fishing to keep her

3
gram

They also point to ninety nine point six billion dollars from two hundred
four oh six three percent of the rates of interest stores as Mexico and
Brazil on market conditions

Taalmodel

Model van de waarschijnlijkheid waarmee woorden op elkaar volgen

- Modelleert de waarschijnlijkheid waarmee woorden op elkaar volgen
- Kan o.a. gebruikt worden om teksten te generen, aan te vullen en te verbeteren
- Markoviaanse assumptie nodig om met productiviteit van taal te kunnen omgaan (max 5-grams)

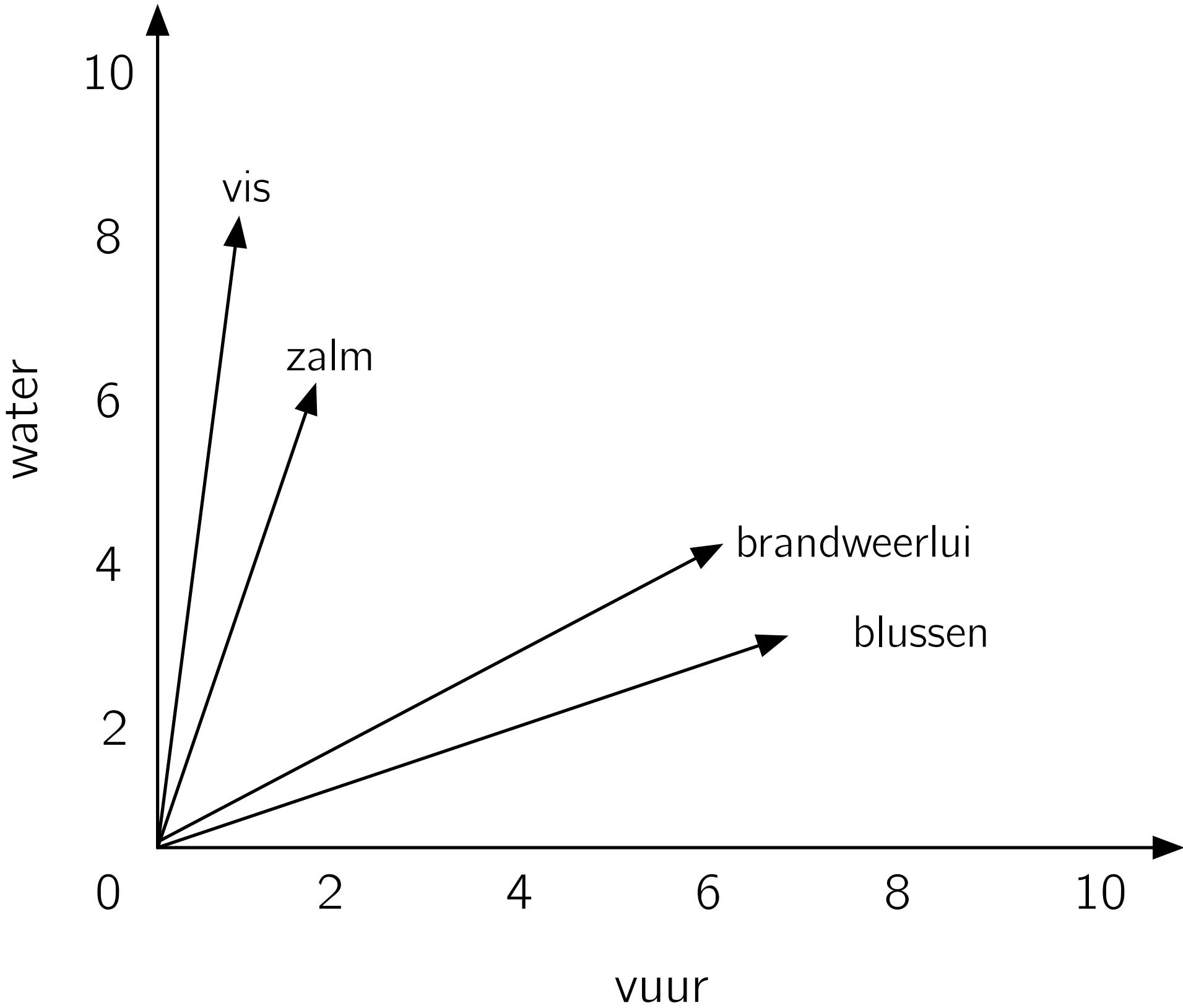
Distributionele semantiek

Voorstelling van woorden in functie van hun waarschijnlijke burens

- De betekenis van een woord kan gemodelleerd worden via de woorden die gewoonlijk in zijn buurt (*context*) voorkomen
- **Distributionele hypothese:** woorden die gewoonlijk in gelijkaardige contexten voorkomen hebben een gelijkaardige betekenis

[...] er zaten best wel veel vissen in het water, vooral zalm [...] de brandweerlui konden het vuur op tijd blussen [...] een lekker zalm boven open vuur gerookt [...] Niet elk vuur kan je met water blussen [...] zalm kan zowel in zoet als zout water overleven [...]

	vuur	water
vis	1	8
zalm	2	6
brandweerlui	6	4
blussen	7	3



$$\cos(\overrightarrow{A}, \overrightarrow{B}) = \frac{\overrightarrow{A} \cdot \overrightarrow{B}}{|\overrightarrow{A}| |\overrightarrow{B}|}$$

$$\overrightarrow{A} \cdot \overrightarrow{B} = \sum_{i=1}^n A_i B_i$$

$$|\overrightarrow{A}| = \sqrt{\sum_{i=1}^n A_i^2}$$

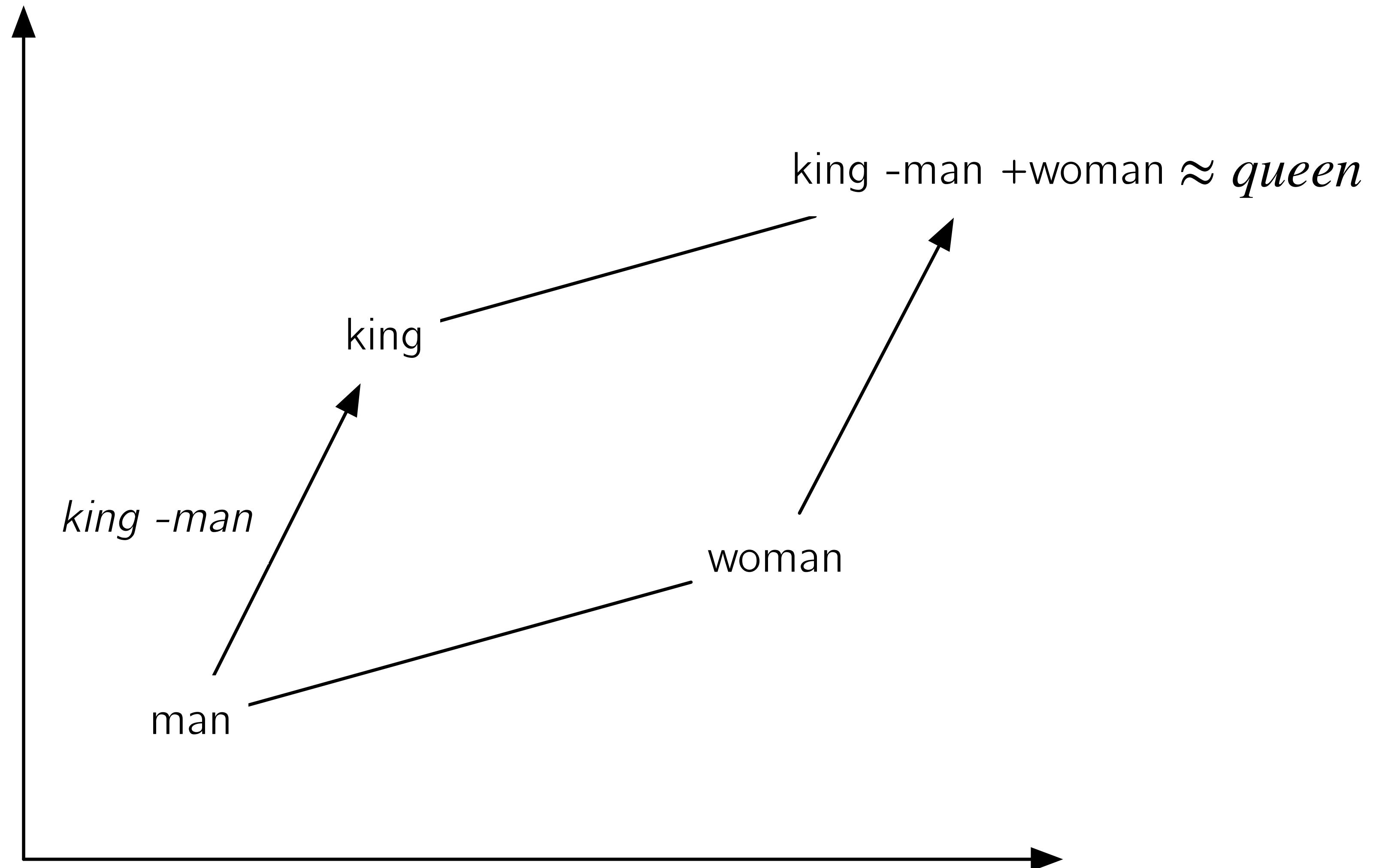
$$\cos(\overrightarrow{vis}, \overrightarrow{zalm}) = \frac{1 * 2 + 8 * 6}{\sqrt{1^2 + 8^2} \sqrt{2^2 + 6^2}} = 0.98$$

$$\cos(\overrightarrow{vis}, \overrightarrow{blussen}) = \frac{1 * 7 + 8 * 3}{\sqrt{1^2 + 8^2} \sqrt{7^2 + 3^2}} = 0.50$$

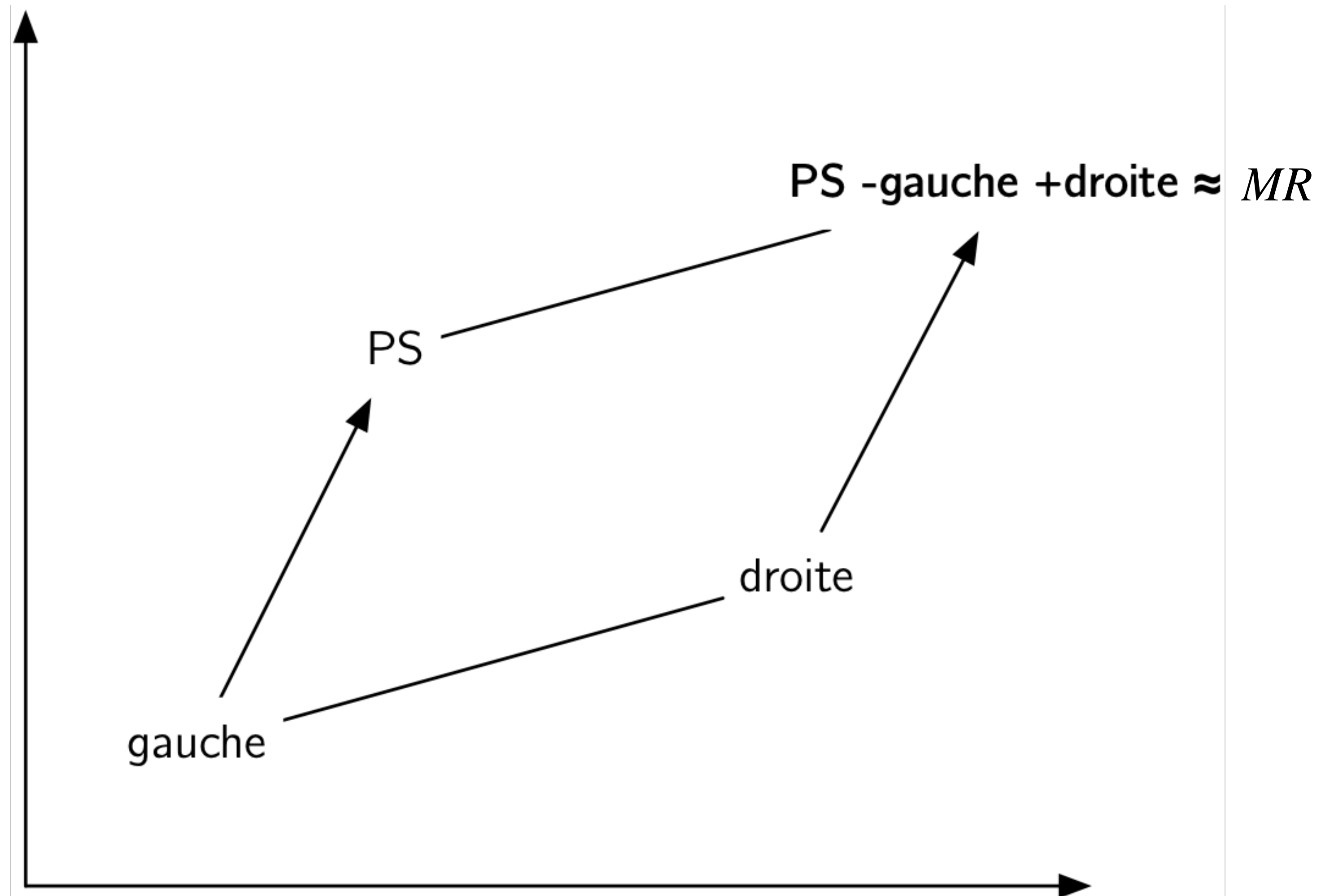
$$\cos(\overrightarrow{brandweer}, \overrightarrow{blussen}) = \frac{6 * 7 + 4 * 3}{\sqrt{6^2 + 4^2} \sqrt{7^2 + 3^2}} = 0.96$$

	aarde	abacus	...	vis	...	machine	...	zoom	zygote
aarde	0	0	...	7	...	0	...	0	0
abacus	0	0	...		...	150	...	0	0
...	...	...	...	...	...	...	...	...	...
vis	7	0	...	0	...	0	...	0	2
...	...	...	...	...	...	...	...	...	...
machine	0	150	...	0	...	0	...	23	0
...	...	...	...	...	...	...	...	...	...
zoom	0	0	...	0	...	23	...	0	0
zygote	0	0	...	2	...	0	...	0	0

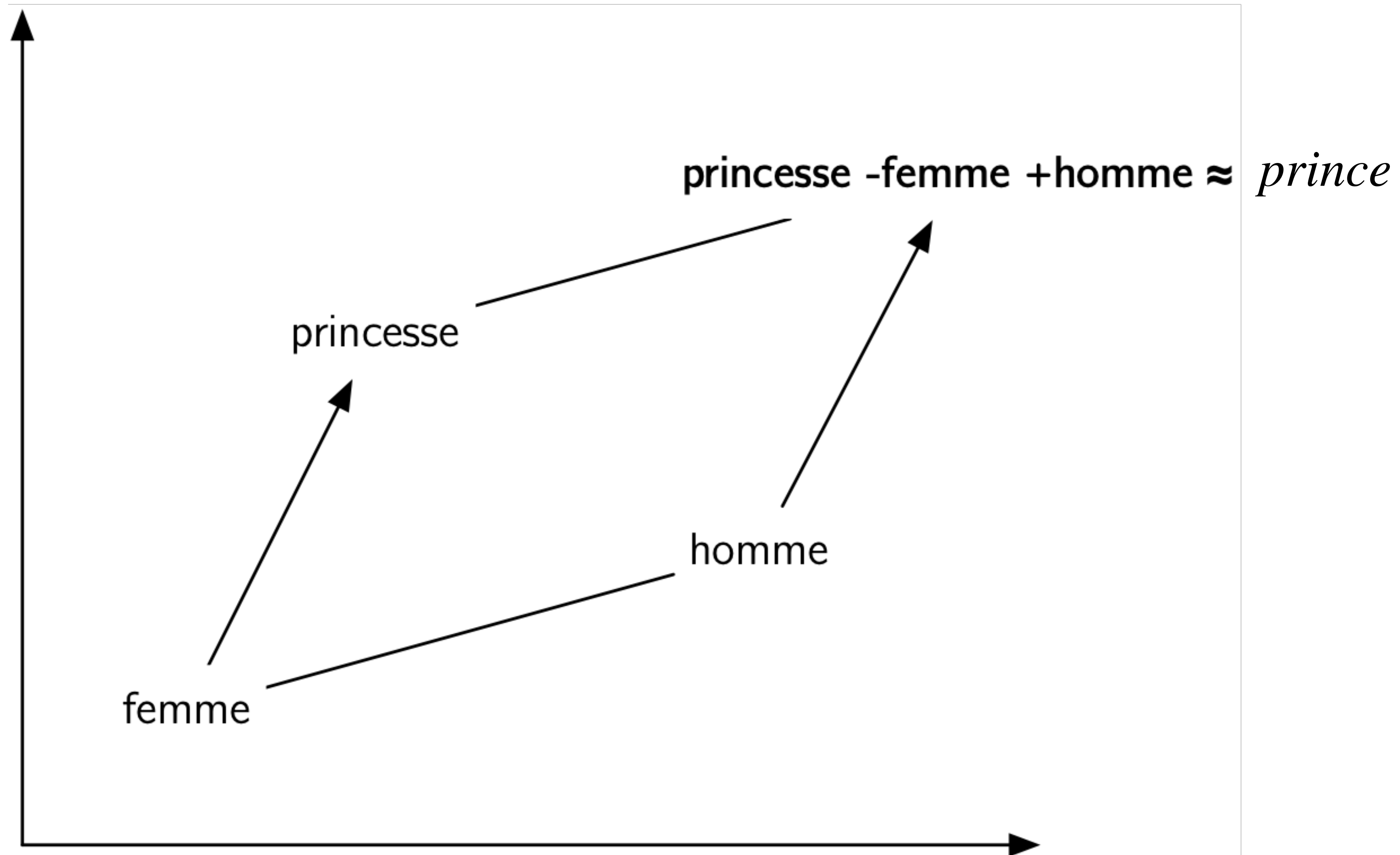
"Redeneren" met woordvectoren



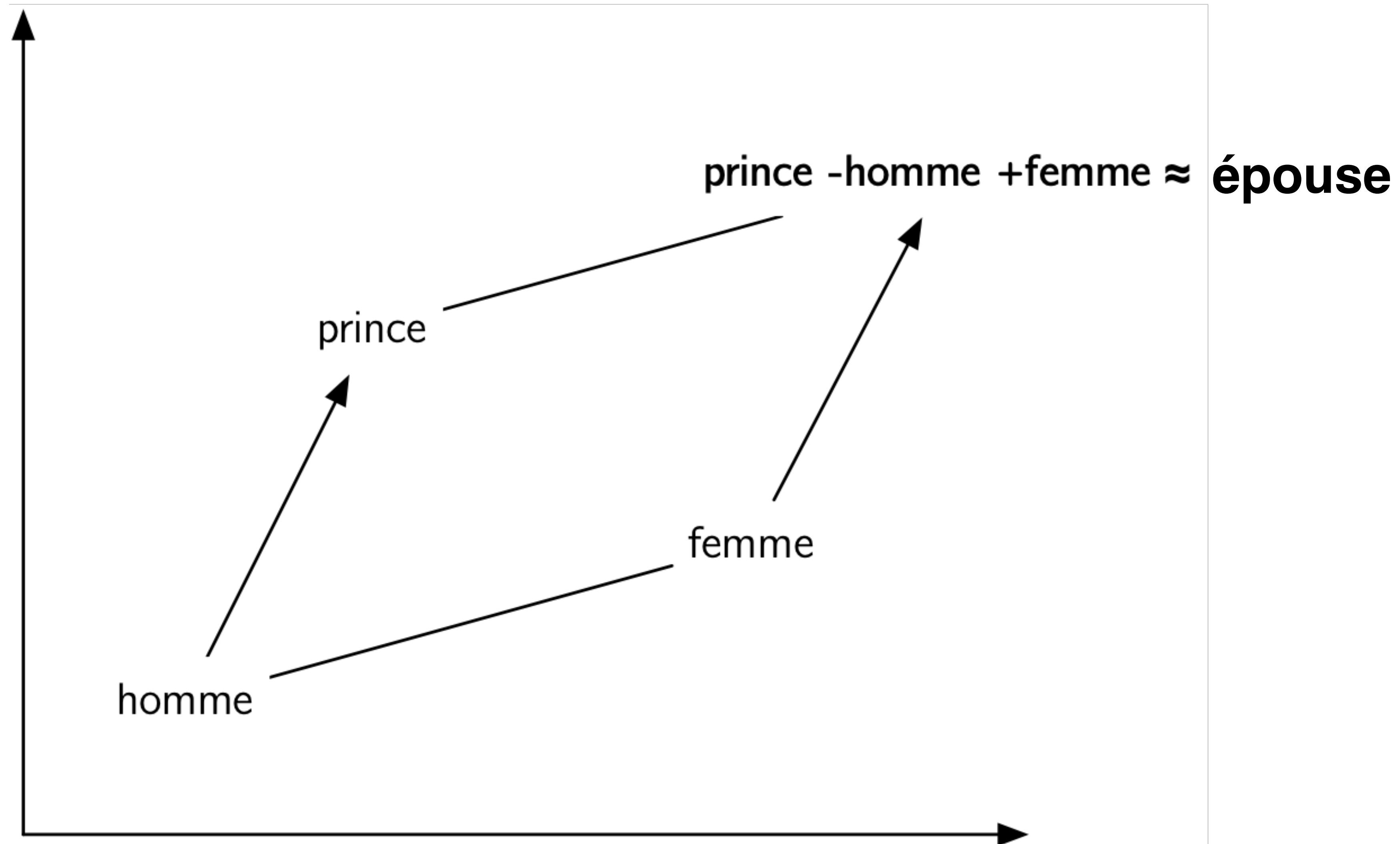
"Redeneren" met woordvectoren



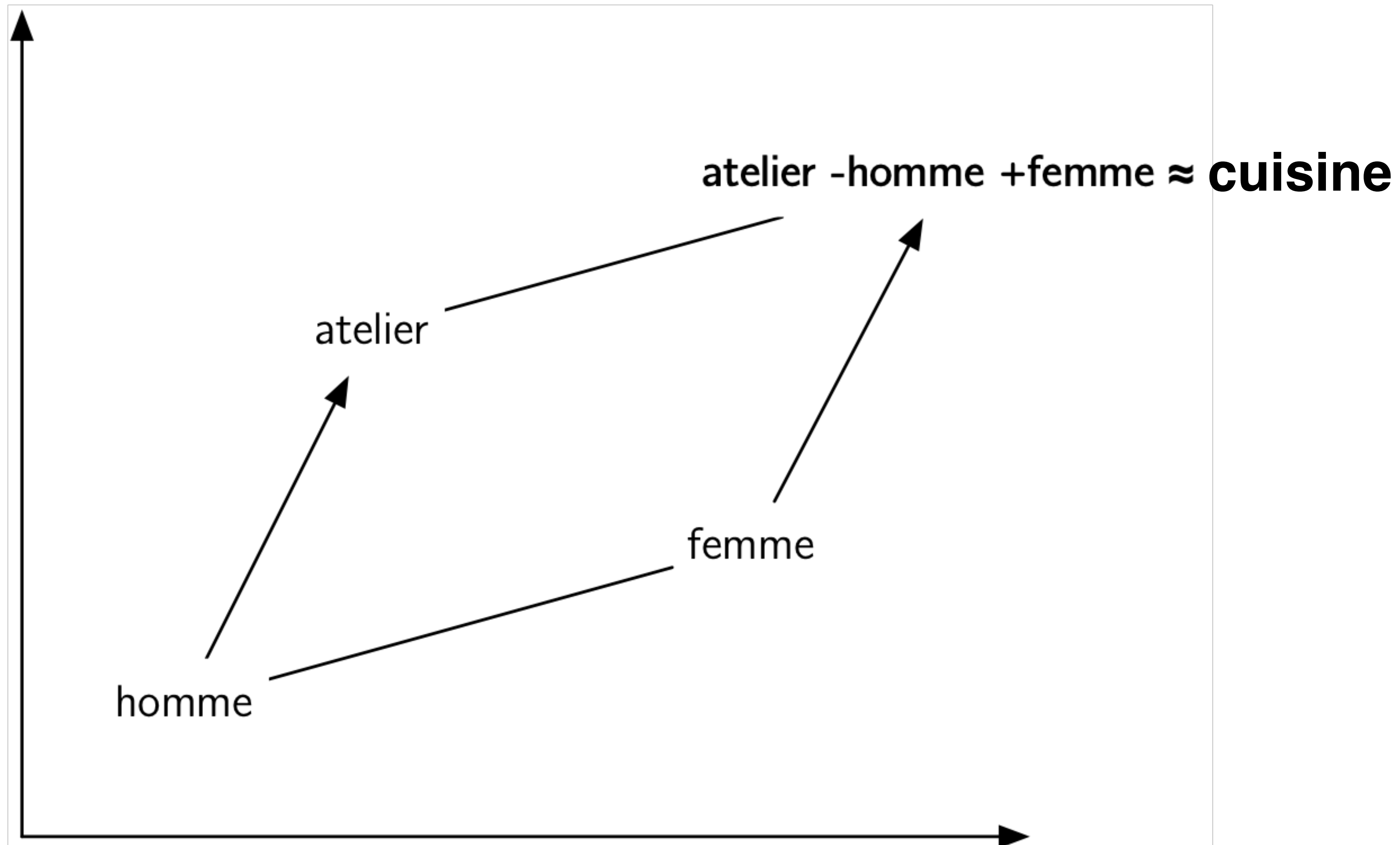
"Redeneren" met woordvectoren



"Redeneren" met woordvectoren



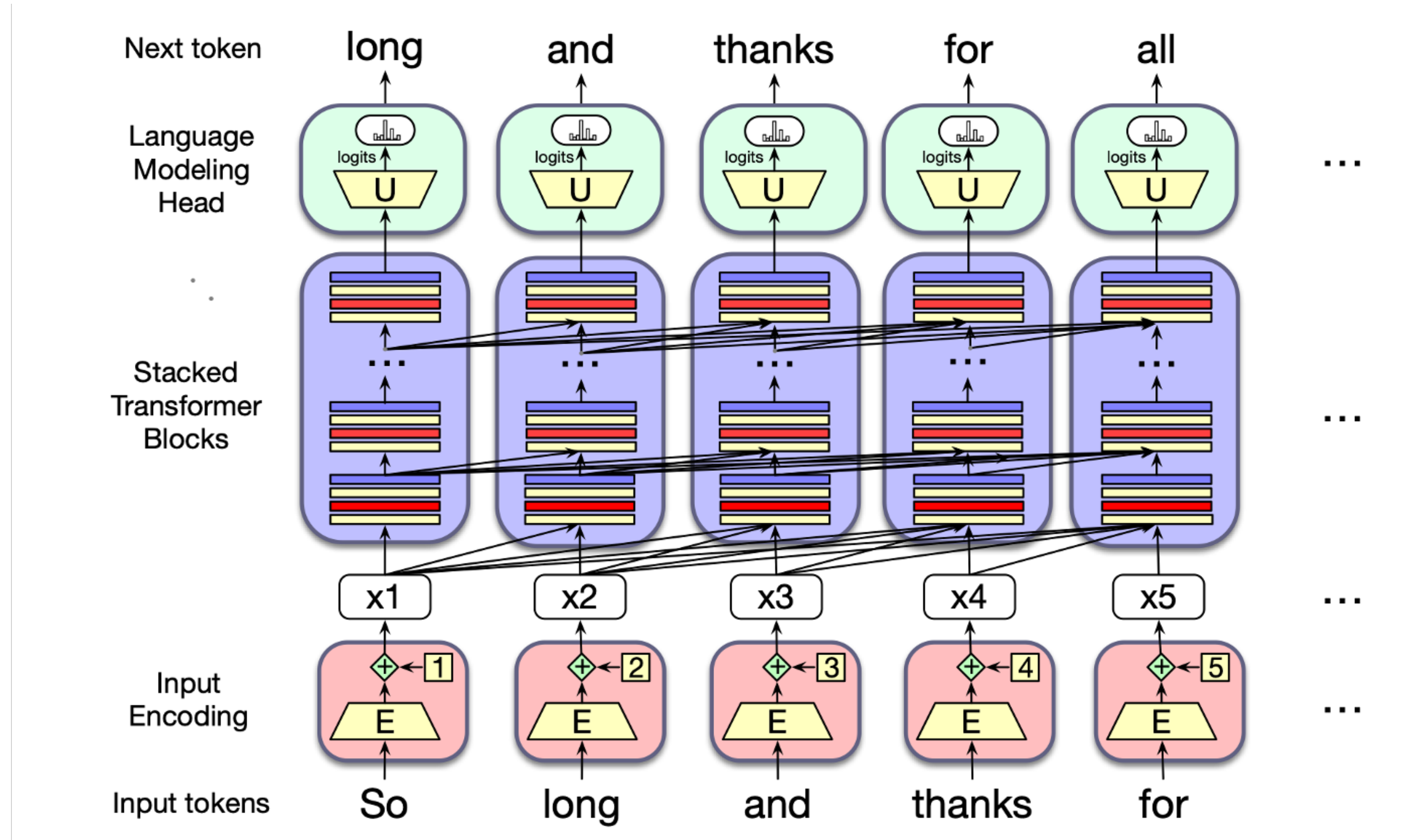
"Redeneren" met woordvectoren



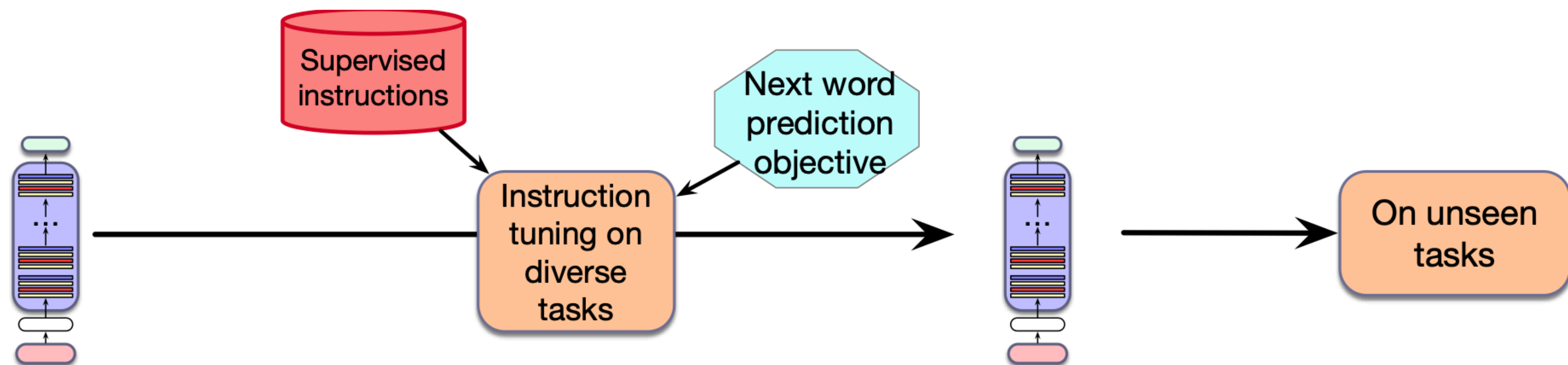
Distributionele semantiek

- Woorden worden voorgesteld als vectoren die hun co-occurentie met andere woorden vatten
- Deze voorstelling ondersteunen een vorm van "redenering"
- Opgelet! De redenering gebeurt puur op basis van de distributie van woorden, NIET op basis van logica of betekenis

Large Language Models



Instruction Tuning (SFT)



Onder de motorkap van ChatGPT

- Generatieve taalmodellen voorspellen teksten op basis van woorddistributies geleerd uit enorme hoeveelheden data
- Grootschalige tuning om te voldoen aan menselijke voorkeuren
- Geen garanties over factuele correctheid, logische correctheid of consistentie

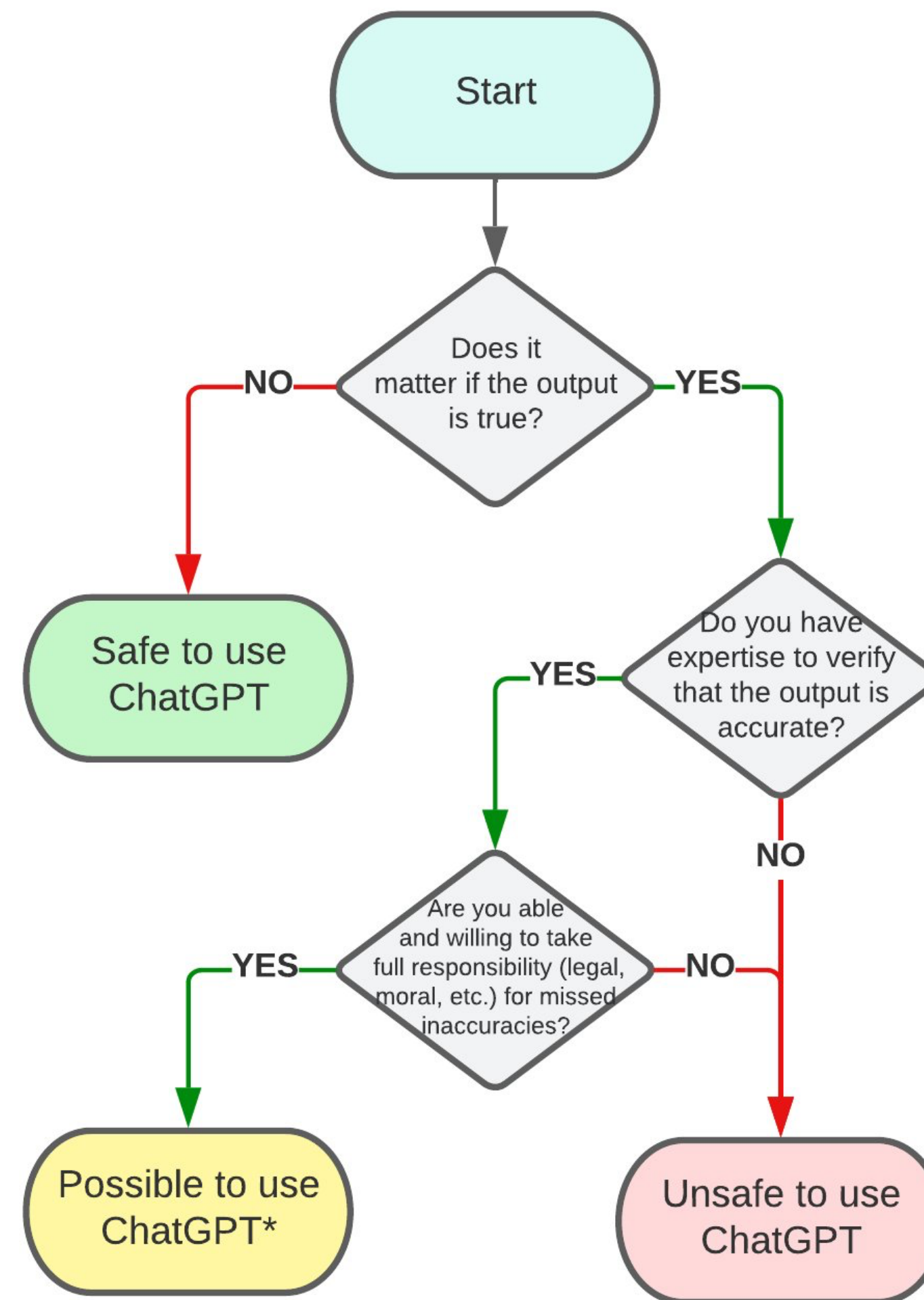
Onder de motorkap van ChatGPT

- Let op voor antropomorfisme en het ELIZA-effect!
 - LLMs *hallucineren* niet, maar het is nu eenmaal zo dat een opeenvolging van woorden die *waarschijnlijk* is niet noodzakelijk ook *waar* is.
 - LLMs *denken* of *redeneren* niet, maar voeren operaties uit in een vectorruimte
- LLMs *denken* of *redeneren* niet, ook al beantwoorden ze redeneerproblemen wel vaak correct.

Vragen hoe slim ChatGPT is, is als vragen hoe sportief mijn stofzuiger is. Het slaat nergens op.

Is it safe to use ChatGPT for your task?

Aleksandr Tiulkanov | January 19, 2023



* but be sure to verify each output word and sentence for accuracy and common sense



Artificiële Intelligentie

Faculteit Wetenschappen en Bio-ingenieurswetenschappen

Heb jij een brede interesse in zowel technologie als menswetenschappen? Studeer Artificiële Intelligentie aan de VUB. Met een unieke mix van psychologie, computerwetenschappen en taalkunde, is deze bacheloropleiding precies wat je zoekt.

Schrijf je nu in!

Introductie

Programma

Iets voor jou?

Na je studie

Waarom VUB?

Toelating & inschrijven