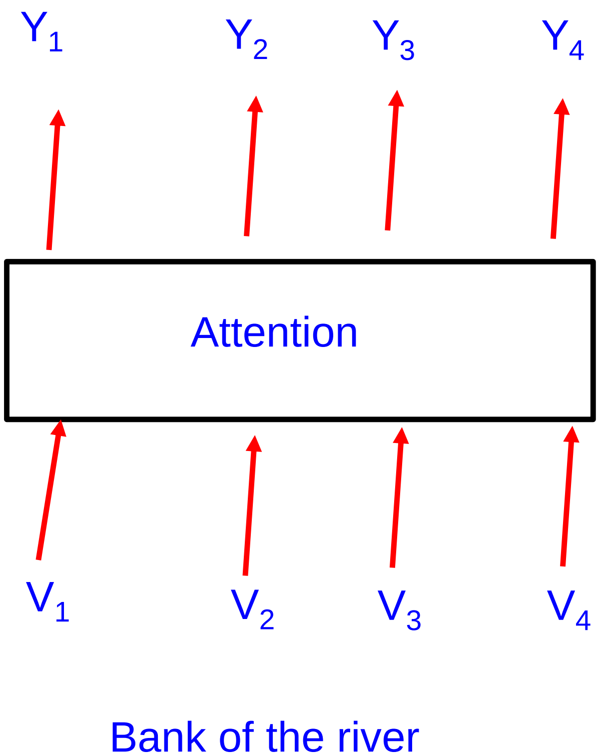
CS772: Deep Learning for Natural Language Processing (DL-NLP)

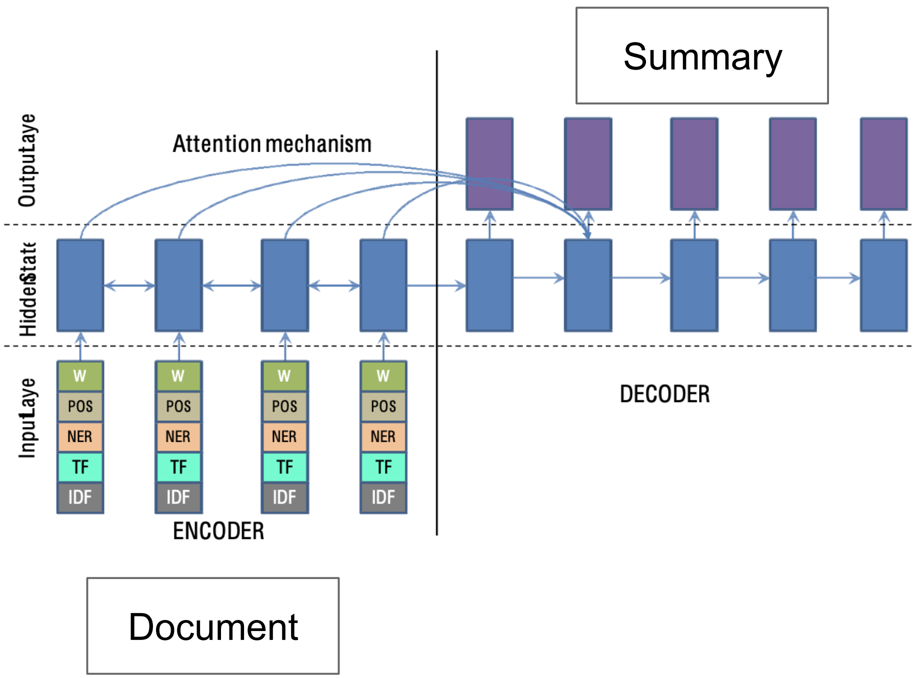
Attention and Alignment

Pushpak Bhattacharyya
Computer Science and Engineering Department
IIT Bombay

Week 12 of 25mar24

1-slide recap



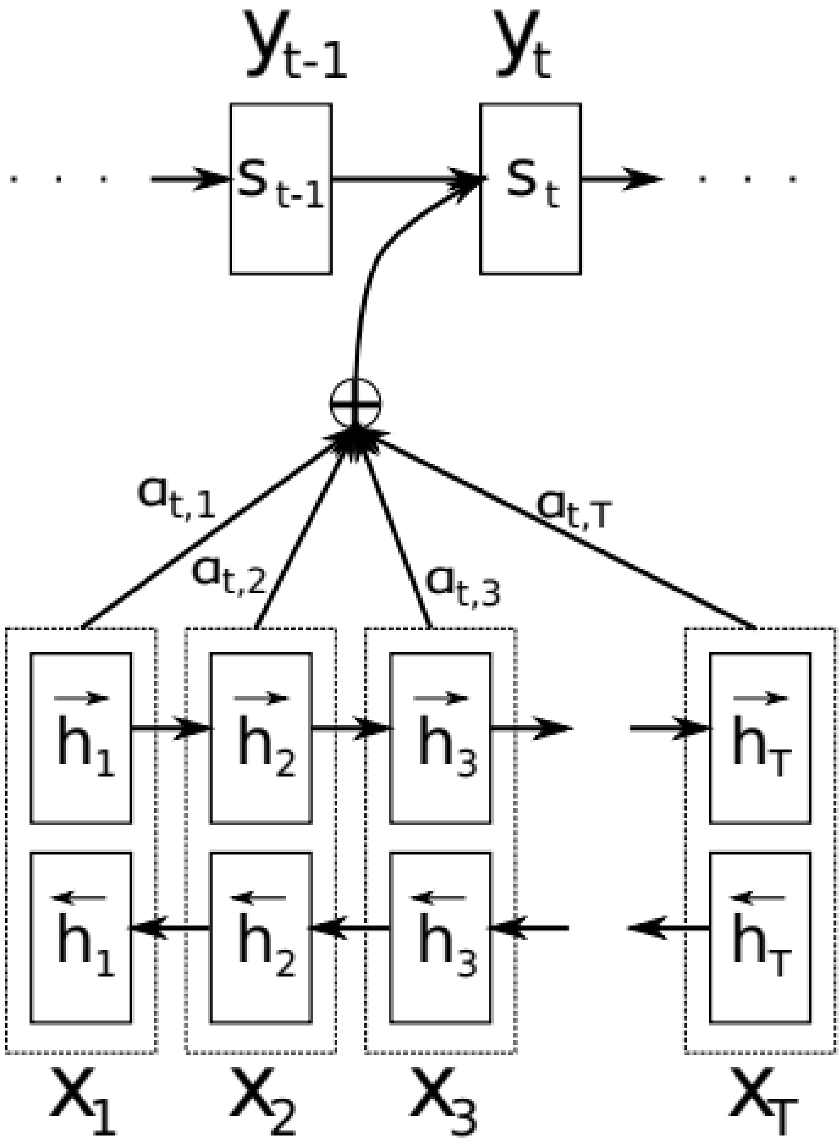


$$Y_i - vector$$
$$= w_{11}.V_1 + w_{12}.V_2 + w_{13}.V_3 + w_{14}.V_4$$

Neural Cross Attention
vs.
Statistical Alignment Learning

Non-neural

Recap of cross attention (Bahadanu et al, 2015)



$$h_t = f(x_t, h_{t-1})$$

$$c = g(\{h_1, h_2, h_3, \ldots h_{T_x}\})$$

$$P(\bar{y}) = \prod_{t=1}^{T} P(y_t \mid \{y_1, y_2, y_3, \ldots y_{t-1}\}, c)$$

$$P(y_t \mid \{y_1, y_2, y_3, \ldots y_{t-1}\}, c) = g(y_{t-1}, s_t, c)$$

$$p(y_i \mid y_1, y_2, \ldots y_{i-1}) = g(y_{i-1}, s_i, c_i)$$

$$s_i = f(s_{i-1}, y_{i-1}, c_i)$$

$$c_i = \sum_{j=1}^{T_x} \alpha_{ij} h_j$$

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^{T_x} \exp(e_{ik})}$$

$$e_{ij} = a(s_{i-1}, h_j)$$

Foundation

- Data driven approach
- Goal is to find out the English sentence *e* given foreign language sentence *f* whose *p(e|f)* is maximum.

$$\tilde{e} = \underset{e \in e^*}{\operatorname{argmax}}\, p(e|f) = \underset{e \in e^*}{\operatorname{argmax}}\, p(f|e)p(e)$$

- Translations are generated on the basis of statistical model
- Parameters are estimated using bilingual parallel corpora

Different kinds of alignment

Correspondence between tokens across translations can be one to one, many to one or one to many e.g.

Ram went to school to play.

राम खेलने के लिए पाठशाला गया

(*raam khelne ke liye paathshaala gayaa)*

Ram <=> राम (one to one)

to play <=> खेलने के लिए (many to many)

to school <=> पाठशाला (many to one)

SMT: Language Model

- To detect *good* English sentences
- Probability of an English sentence $w_1w_2 \ldots\ldots w_n$ can be written as $Pr(w_1w_2 \ldots\ldots w_n) = Pr(w_1) * Pr(w_2|w_1) * \ldots * Pr(w_n|w_1\, w_2 \ldots w_{n-1})$
- Here $Pr(w_n|w_1\, w_2 \ldots w_{n-1})$ is the probability that word w_n follows word string $w_1\, w_2 \ldots w_{n-1}$.
  - N-gram model probability
- Trigram probability calculation

$$p(w_3|w_1w_2) = \frac{count(w_1w_2w_3)}{count(w_1w_2\,)}$$

SMT: Translation Model

- $P(f|e)$: Probability of some f given hypothesis English translation e
- How to assign the values to $p(e|f)$? Sentences are infinite, not possible to find pair (e,f) for all sentences

$$p(f|e) = \frac{count(f,e)}{count(e)} \longleftarrow$$ Sentence level

- Introduce a hidden variable ***a***, that represents alignments between the individual words in the sentence pair

$$\Pr(\boldsymbol{f}|\boldsymbol{e}) = \sum_{\boldsymbol{a}} \Pr(\boldsymbol{f}, \boldsymbol{a}|\boldsymbol{e}) \longleftarrow$$ Word level

Alignment

- If the string, $e = e_1^l = e_1\ e_2\ \ldots e_l$, has l words, and the string, $f = f_1^m = f_1 f_2 \ldots f_m$, has m words,
- then the alignment, a, can be represented by a series, $\boldsymbol{a_1^m = a_1 a_2 \ldots a_m}$, of m values, each between 0 and l such that if the word in position j of the f-string is connected to the word in position i of the e-string, then
  - $\boldsymbol{a_j = i}$, and
  - if it is not connected to any English word, then $\boldsymbol{a_j = O}$

Example of alignment

English: *Ram went to school*

Hindi: *raam paathashaalaa gayaa*

Ram *went* *to* *school*

<Null> *raam* *paathashaalaa* *gayaa*

Translation Model: Exact expression

$$\Pr(\boldsymbol{f}, \boldsymbol{a}|\boldsymbol{e}) = \Pr(m|\boldsymbol{e}) \prod_{j=1}^{m} \Pr\left(a_j \,|a_1^{j-1}, f_1^{j-1}, m, \boldsymbol{e}\right) \Pr\left(f_j | a_1^{j}, f_1^{j-1}, m, \boldsymbol{e}\right)$$

Choose the length of foreign language string given *e*

Choose alignment given *e* and *m*

Choose the identity of foreign word given *e, m, a*

- Five models for estimating parameters in the expression [2]
- Model-1, Model-2, Model-3, Model-4, Model-5

Proof of Translation Model: Exact expression

$$\Pr(f \mid e) = \sum_a \Pr(f, a \mid e)$$; marginalization

$$\Pr(f, a \mid e) = \sum_m \Pr(f, a, m \mid e)$$; marginalization

$$\begin{aligned}
\sum_m \Pr(f, a, m \mid e) &= \sum_m \Pr(m \mid e) \Pr(f, a \mid m, e) \\
&= \sum_m \Pr(m \mid e) \Pr(f, a \mid m, e) \\
&= \sum_m \Pr(m \mid e) \prod_{j=1}^{m} \Pr(f_j, a_j \mid a_1^{j-1}, f_1^{j-1}, m, e) \\
&= \sum_m \Pr(m \mid e) \prod_{j=1}^{m} \Pr(a_j \mid a_1^{j-1}, f_1^{j-1}, m, e) \Pr(f_j \mid a_1^{j}, f_1^{j-1}, m, e)
\end{aligned}$$

m is fixed for a particular *f*, hence

$$\sum_m \Pr(f, a, m \mid e) = \Pr(m \mid e) \prod_{j=1}^{m} \Pr(a_j \mid a_1^{j-1}, f_1^{j-1}, m, e) \Pr(f_j \mid a_1^{j}, f_1^{j-1}, m, e)$$

$$\Pr(f \mid e) = \Pr(m \mid e) . \sum_a \left[\prod_{j=1}^{m} \Pr(a_j \mid a_1^{j-1}, f_1^{j-1}, m, e) \Pr(f_j \mid a_1^{j}, f_1^{j-1}, m, e) \right]$$

EM for word alignment from sentence alignment: example

English	French
(1) three rabbits	(1) trois lapins
a b	w x
(2) rabbits of Grenoble	(2) lapins de Grenoble
b c d	x y z

Initial Probabilities:
each cell denotes *t(a←→ w), t(a←→ x) etc.*

	a	b	c	d
w	1/4	1/4	1/4	1/4
x	1/4	1/4	1/4	1/4
y	1/4	1/4	1/4	1/4
z	1/4	1/4	1/4	1/4

Example of expected count

$C[w \leftarrow\rightarrow a;\ (a\ b) \leftarrow\rightarrow (w\ x)]$

$$= \frac{t(w \leftarrow\rightarrow a)}{t(w \leftarrow\rightarrow a) + t(w \leftarrow\rightarrow b)} \times \#(a \text{ in } `a\ b\text{'}) \times \#(w \text{ in } `w\ x\text{'})$$

$$= \frac{1/4}{1/4+1/4} \times 1 \times 1 = 1/2$$

"counts"

a b ←→ ***w x***	a	b	c	d
w	1/2	1/2	0	0
x	1/2	1/2	0	0
y	0	0	0	0
z	0	0	0	0

b c d ←→ ***x y z***	a	b	c	d
w	0	0	0	0
x	0	1/3	1/3	1/3
y	0	1/3	1/3	1/3
z	0	1/3	1/3	1/3

Revised probability: example

$$t_{revised}(a \leftarrow\rightarrow w)$$

$$= \frac{1/2}{(1/2+1/2+0+0)_{(a\ b)\leftarrow\rightarrow(w\ x)} + (0+0+0+0)_{(b\ c\ d)\leftarrow\rightarrow(x\ y\ z)}}$$

Revised probabilities table

	a	b	c	d
w	1/2	1/2	0	0
x	1/4	5/12	1/6	1/6
y	0	1/3	1/3	1/3
z	0	1/3	1/3	1/3

“revised counts”

a b ←→ w x	a	b	c	d
w	1/2	3/8	0	0
x	1/2	5/8	0	0
y	0	0	0	0
z	0	0	0	0

b c d ←→ x y z	a	b	c	d
w	0	0	0	0
x	0	5/9	1/3	1/3
y	0	2/9	1/3	1/3
z	0	2/9	1/3	1/3

Re-Revised probabilities table

	a	b	c	d
w	1/2	1/2	0	0
x	3/16	**85/144**	1/9	1/9
y	0	1/3	1/3	1/3
z	0	1/3	1/3	1/3

Continue until convergence; notice that (b,x) binding gets progressively stronger;
b=rabbits, x=lapins

RLHF

Presented by Swaroop

Reinforcement Learning in Natural Language Generation

Swaroop Nath

Why Reinforcement Learning? | Key Insight

Token-by-Token match

$$\mathcal{L}_{\mathcal{MLE}} = -\sum_{t=1}^{n} \left\{ \boxed{\mathbb{1}[y_t = y_t^*]} \log \mathcal{P}(y_t^* | y_1^*, y_2^*, \cdots, y_{t-1}^*, \mathbf{q}, \mathbf{d}) \right\}$$

$$\mathcal{L}_{\mathcal{PG}} = -\sum_{t=1}^{n} \left\{ \boxed{(R(\tau) - b)} \log \mathcal{P}(y_t^* | y_1^*, y_2^*, \cdots, y_{t-1}^*, \mathbf{q}, \mathbf{d}) \right\}$$

Any Generic Reward

Leveraging Domain Knowledge for Efficient Reward Modelling in RLHF

Leveraging Domain Knowledge for Efficient **Reward Modelling** in RLHF

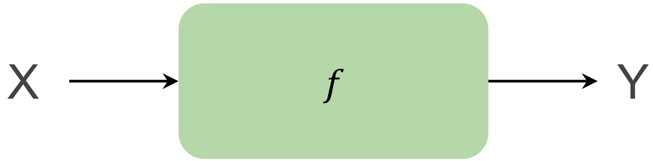
Leveraging Domain Knowledge for Efficient Reward Modelling in **RLHF**

Leveraging **Domain Knowledge** for Efficient Reward Modelling in RLHF

Leveraging Domain Knowledge for Efficient Reward Modelling in RLHF

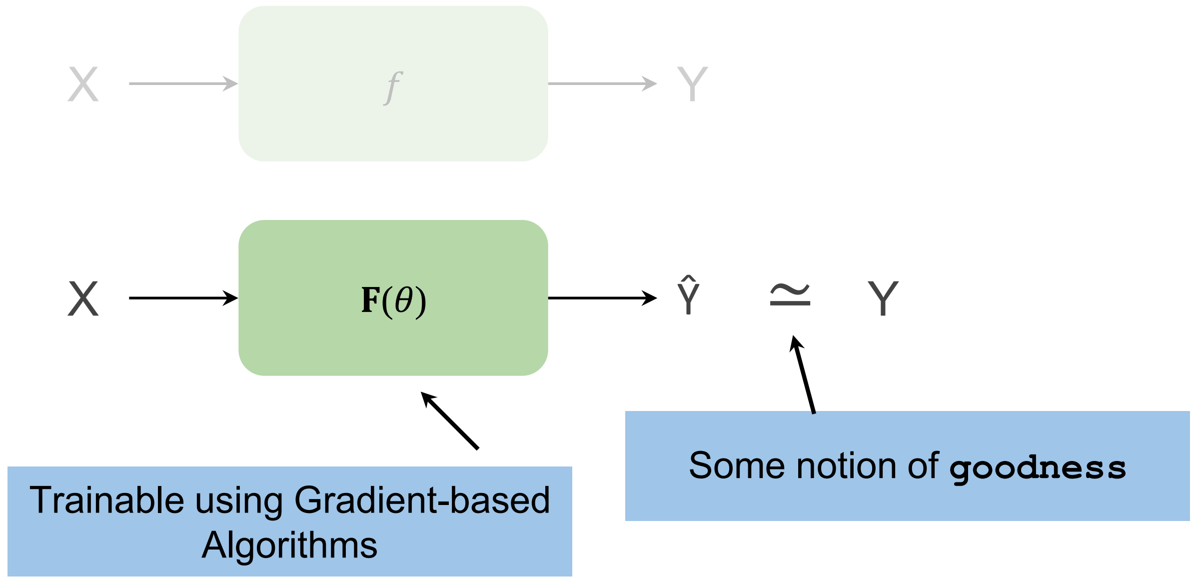
Learning Algorithms

Given inputs (X) and outputs (Y), how do you learn the mapping?



Learning Algorithms

Given inputs (X) and outputs (Y), how do you learn the mapping?



What is this notion of goodness?

Text Generation

The Department of Computer Science and Engineering at IIT Bombay is thrilled to announce the upcoming RISC 2024 . . .

The Computer Science of and Department Engineering upcoming at is thrilled to the announce RISC 2024 IIT Bombay . . .

Fluency

What is this notion of goodness?

Text Generation

Indian Institute of Technology Bombay is in India.

Fluency

Indian Institute of Technology Bombay is in the Australian continent, located in New Zealand.

Hallucination

What is this notion of goodness?

Text Generation

Indian Institute of Technology Bombay is in India.

L'Indian Institute of Technology Bombay se trouve en Inde.

Fluency

Hallucination

Language

What is this notion of goodness?

Text Generation

Inclusivity	Tonality	Clarity	Fluency
Bias and Stereotypes	Originality	Coherence	Hallucination
Audience Appropriateness	Engagement	Conciseness	Language

What is this notion of goodness?

Text Generation

Inclusivity

Bias and Stereotypes

Audience Appropriateness

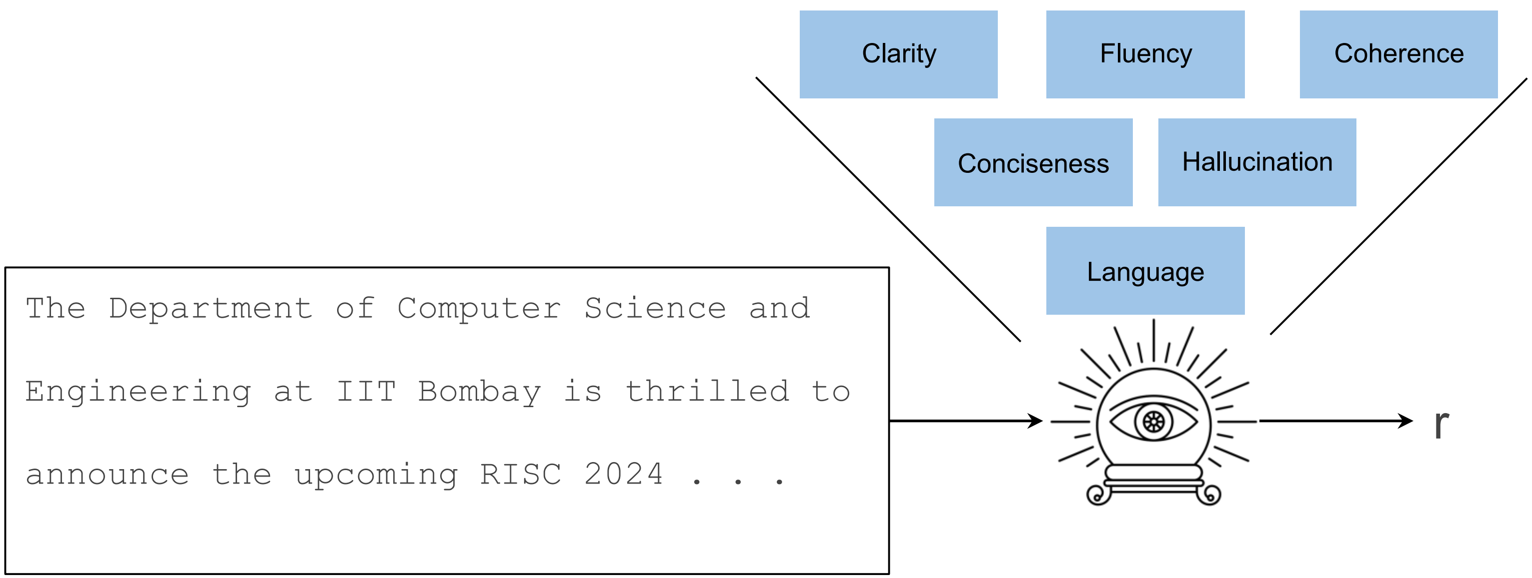
Fluency

Hallucination

Language

An Oracle

Oracle takes all factors into account and scores the text



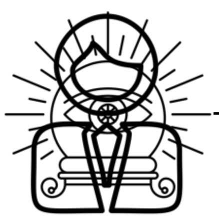
An Oracle

Typically humans play the role of Oracle!

```
The Department of Computer Science and

Engineering at IIT Bombay is thrilled to

announce the upcoming RISC 2024 . . .
```



r

Preference Modeling

Difficult for human to provide a scalar score!

However the score reflects in preference!

What is Preference Feedback?

Another Large Model

Trained on Human Preference Data

What is Preference Feedback?

Another Large Model

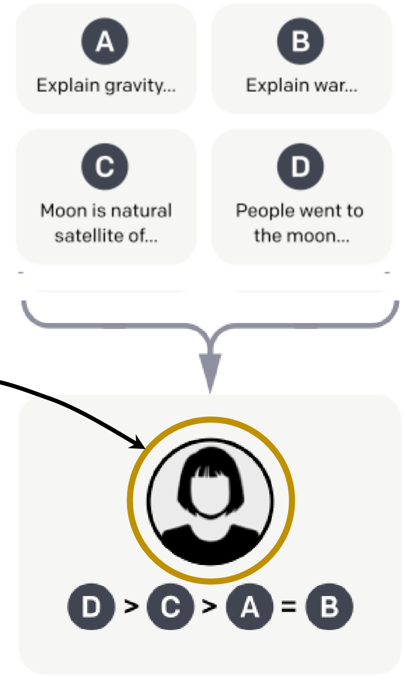
Trained on Human Preference Data

Various Possible Outputs

What is Preference Feedback?

Another Large Model

Trained on Human Preference Data

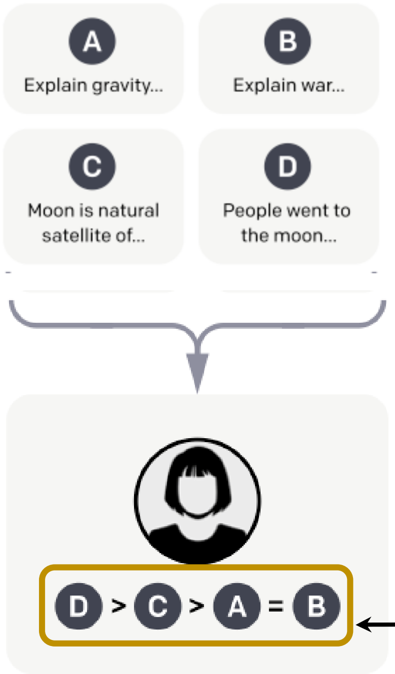


AI Trainer

What is Preference Feedback?

Another Large Model

Trained on Human Preference Data

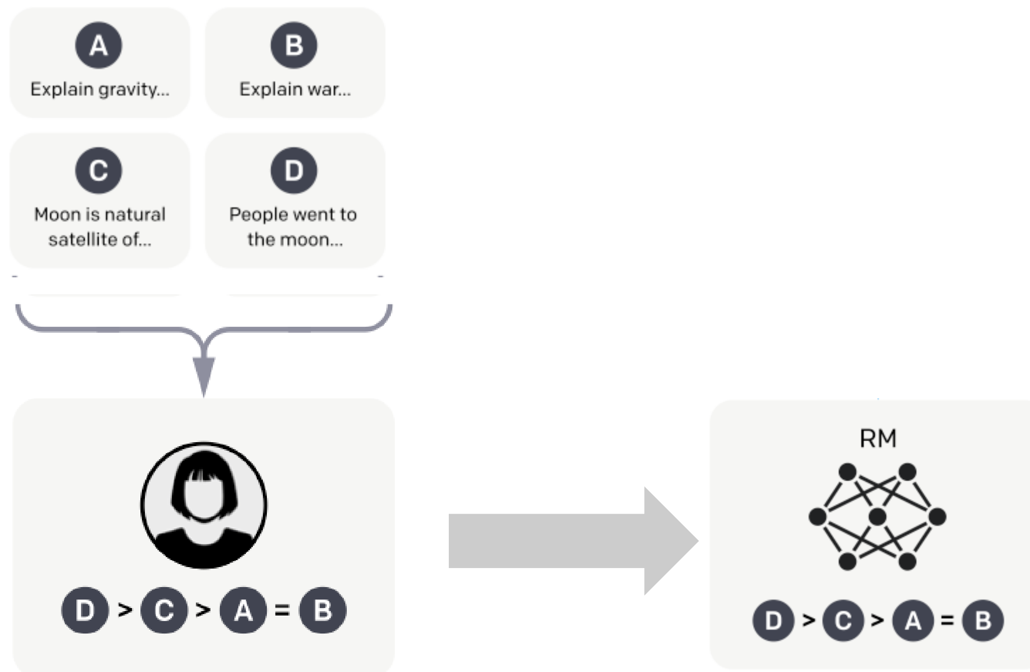


AI Trainer ranks outputs

What is Preference Feedback?

Another Large Model

Trained on Human Preference Data



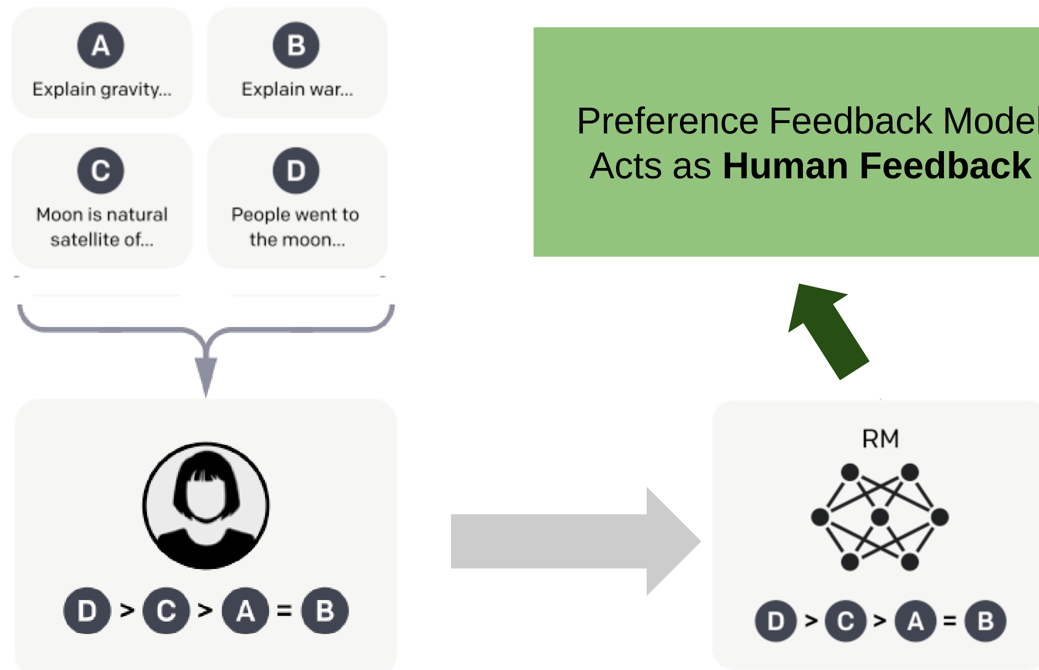
Train Preference Reward Model on this data

What is Preference Feedback?

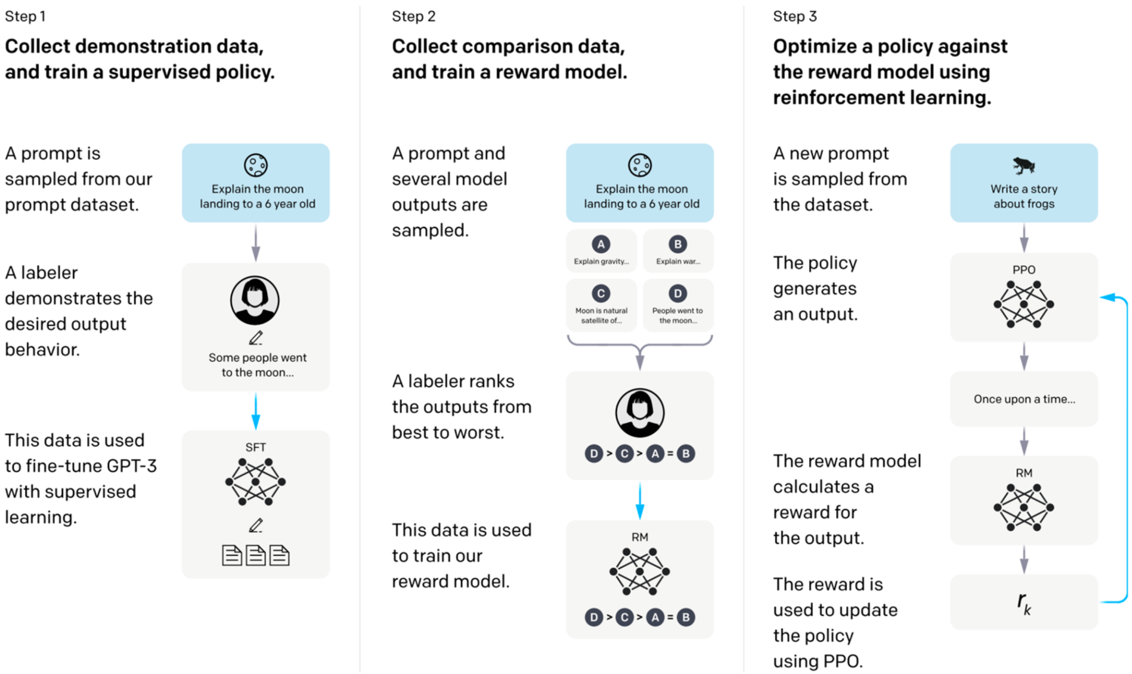
Another Large Model

Trained on Human Preference Data

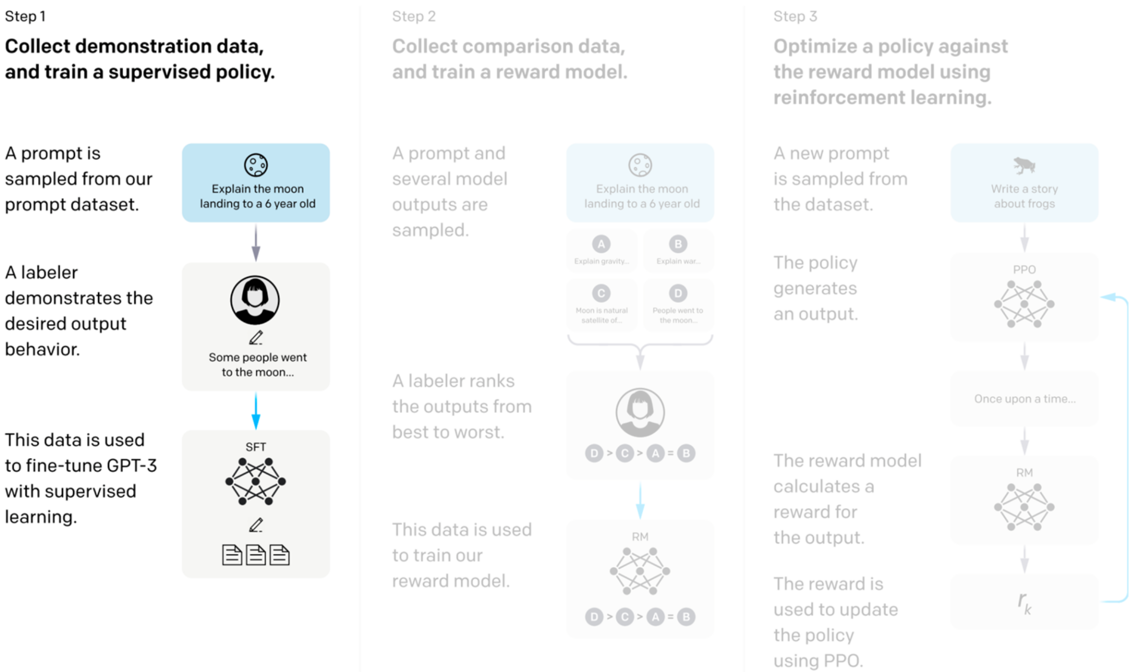
Preference Feedback Model
Acts as **Human Feedback**



Realization: InstructGPT



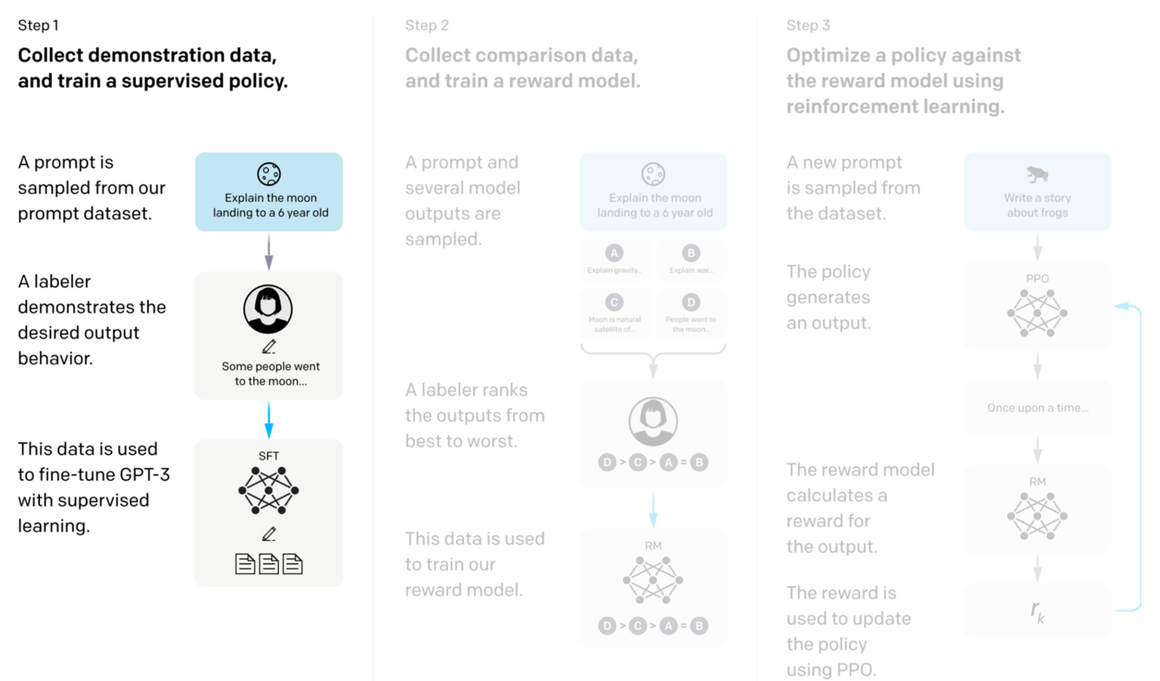
Realization: InstructGPT



Supervised Fine

Tuning

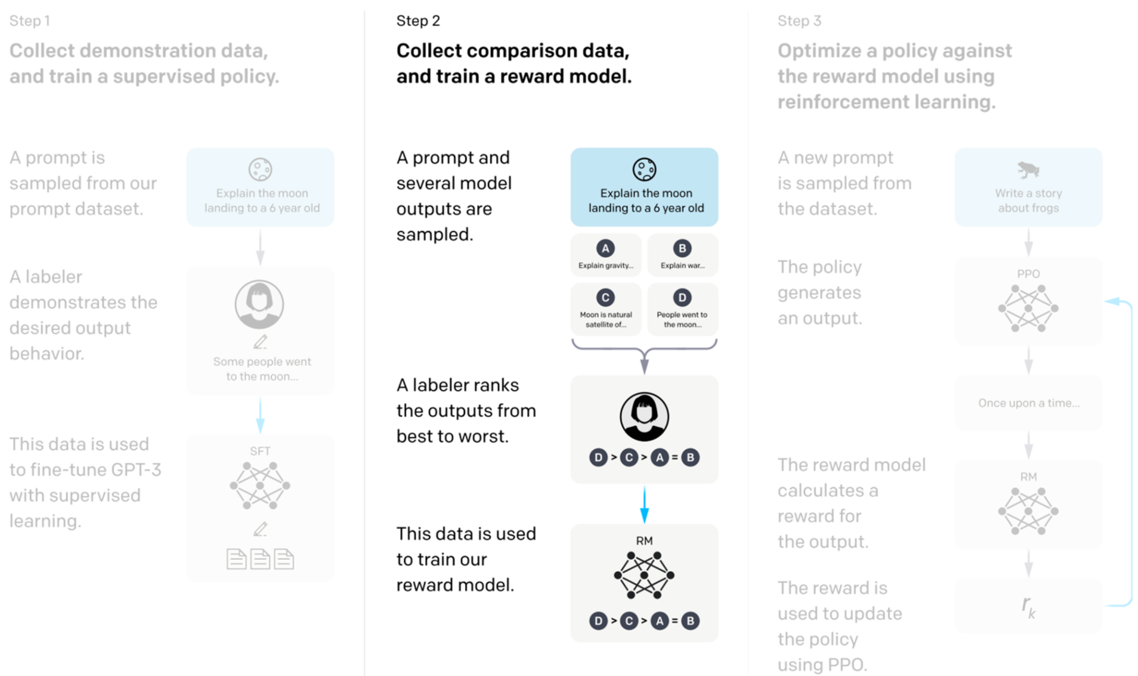
Realization: InstructGPT



Supervised Fine/Instruction Tuning

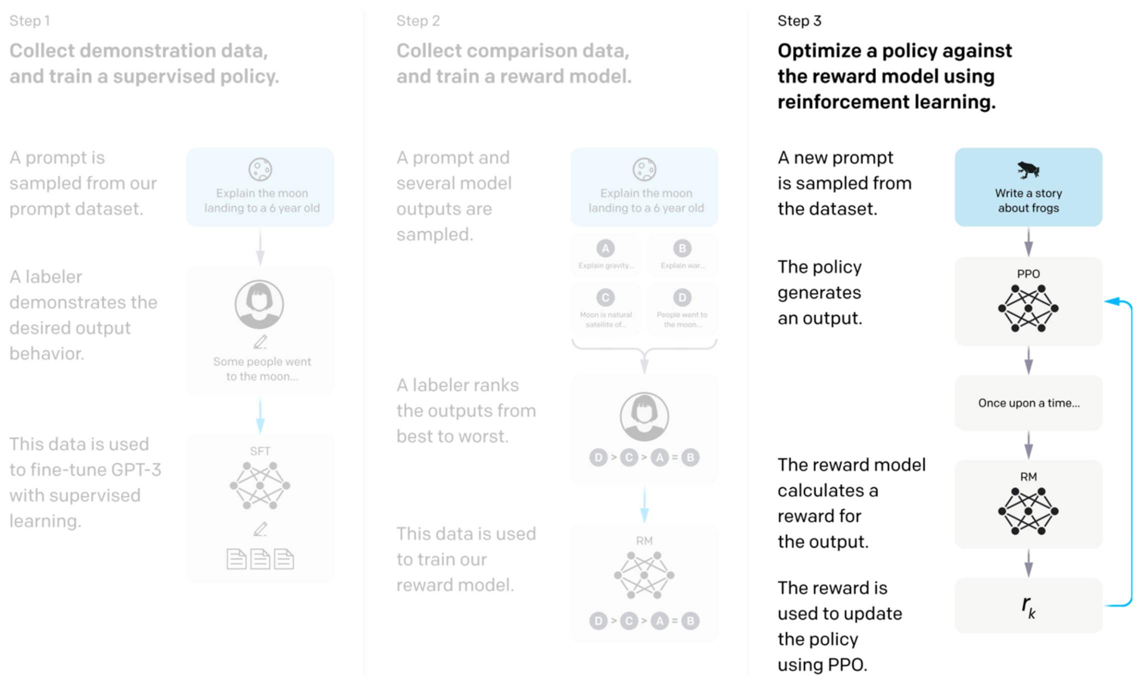
Subtle Difference

Realization: InstructGPT



Collecting data from model for AI Trainer labelling and Preference Model Training

Realization: InstructGPT



RLHF Training using Preference Model to reward generations

Preference Modeling | Underlying Math

Difficult for human to provide a scalar score!

However the score reflects in preference!

Bradley-Terry model[1] provides a way to extract latent-reward-model from preference data

$$p^*(y_1 \succ y_2 \mid x) = \frac{\exp\left(r^*(x, y_1)\right)}{\exp\left(r^*(x, y_1)\right) + \exp\left(r^*(x, y_2)\right)}$$

[1] R. A. Bradley and M. E. Terry. Rank analysis of incomplete block designs: I. the method of paired comparisons. *Biometrika*, 39(3/4):324–345, 1952. doi: https://doi.org/10.2307/2334029

Preference Modeling | Underlying Math

Difficult for human to provide a scalar score!

However the score reflects in preference!

Bradley-Terry model[1] provides a way to extract latent-reward-model from preference data

$$p^*(y_1 \succ y_2 \mid x) = \frac{\exp\left(r^*(x, y_1)\right)}{\exp\left(r^*(x, y_1)\right) + \exp\left(r^*(x, y_2)\right)}$$

[1] R. A. Bradley and M. E. Terry. Rank analysis of incomplete block designs: I. the method of paired comparisons. *Biometrika*, 39(3/4):324–345, 1952. doi: https://doi.org/10.2307/2334029

Requires a lot of human-preference data

Reward Model

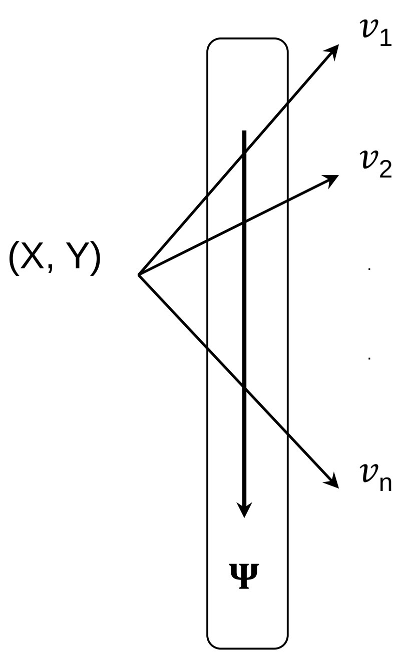
X → input text

Y → output from LLM

$$\varphi = f(X, Y)$$

Reward Model

$\varphi = f(X, Y)$



Hallucination

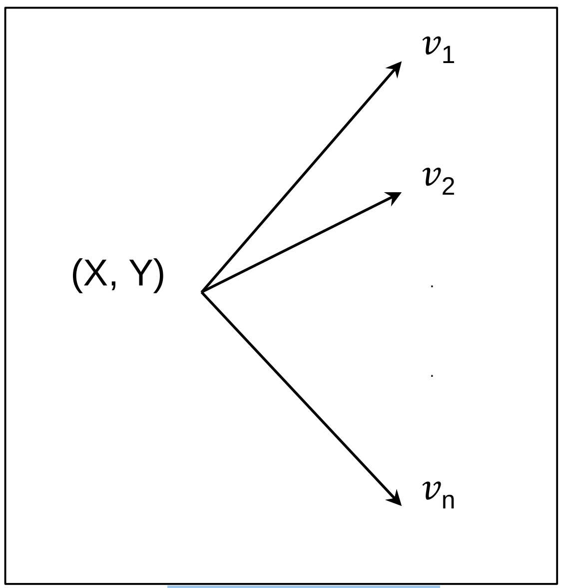
For the task of creative writing

v_i's are specified by domain experts

Signifies what features are important in the reward

Reward Model

$$\varphi = f\,(X, Y)$$



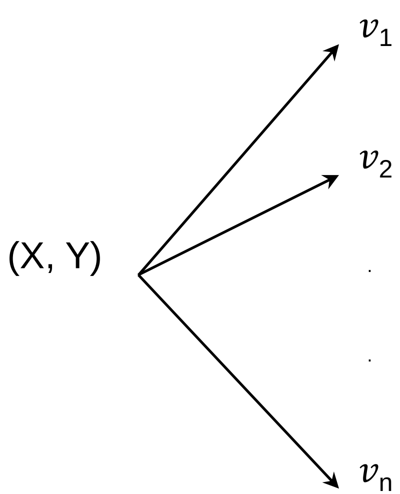
Requires no training data

$$\boxed{\varphi = f\,(v)}$$

Requires training data

Reward Model

$\varphi = f\,(X, Y)$



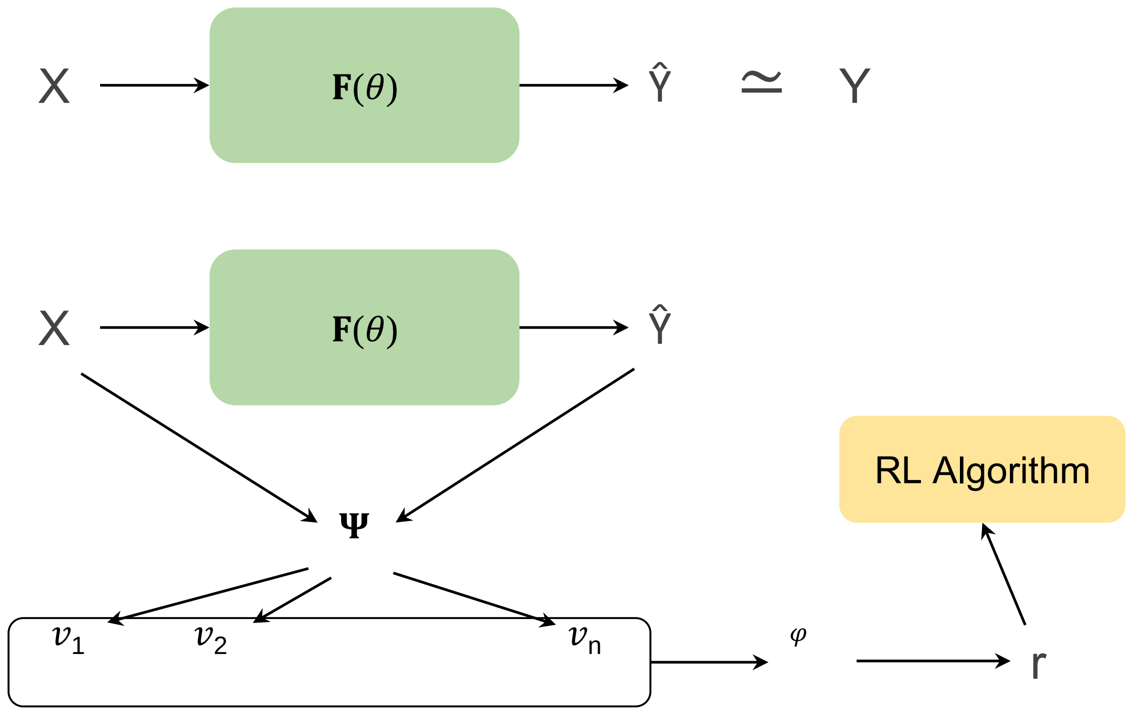
$\varphi = f\,(v)$

Domain Knowledge gives me features

These features compress variability in data

Thus effectively reducing the data requirement

Putting it all together -- RLHF



Ok, but how does it perform on a domain?

E-Commerce Opinion Summarization

Input: Reviews

Output: Opinion Summary

Allen Pinto

Clearness and beautiful performance

Reviewed in India on 5 January 2024

Size: 75 Hz | Style Name: 21.5-Inch Full HD | Verified Purchase

It's HDMI facility and its clarity which helps decipher the print the screen.

Helpful | Report

Sanwar Mal

Immersive Excellence: Unveiling the HP M24f 23.8-Inch Full HD Monitor

Reviewed in India on 24 December 2023

Size: 75 Hz | Style Name: 23.8-Inch Full HD | Verified Purchase

The HP M24f 23.8-Inch Full HD Monitor delivers a visually stunning experience with vibrant colors and sharp resolution. Its sleek design and thin bezels add a touch of sophistication, making it a versatile choice for work or play. With user-friendly features and AMD FreeSync technology for gaming, this monitor combines performance and style seamlessly.

Helpful | Report

Customers say

Customers like the quality, color, value, speed and appearance of the monitor. They mention that its a very good product, works great for Coding, Gaming and Casual Movie watching and that its pretty good for the price value. They are also satisfied with the thickness, and picture quality. That said, opinions are mixed on brightness.

AI-generated from the text of customer reviews

Domain Specification

E-Commerce Opinion Summarization

Input: Reviews

Output: Opinion Summary

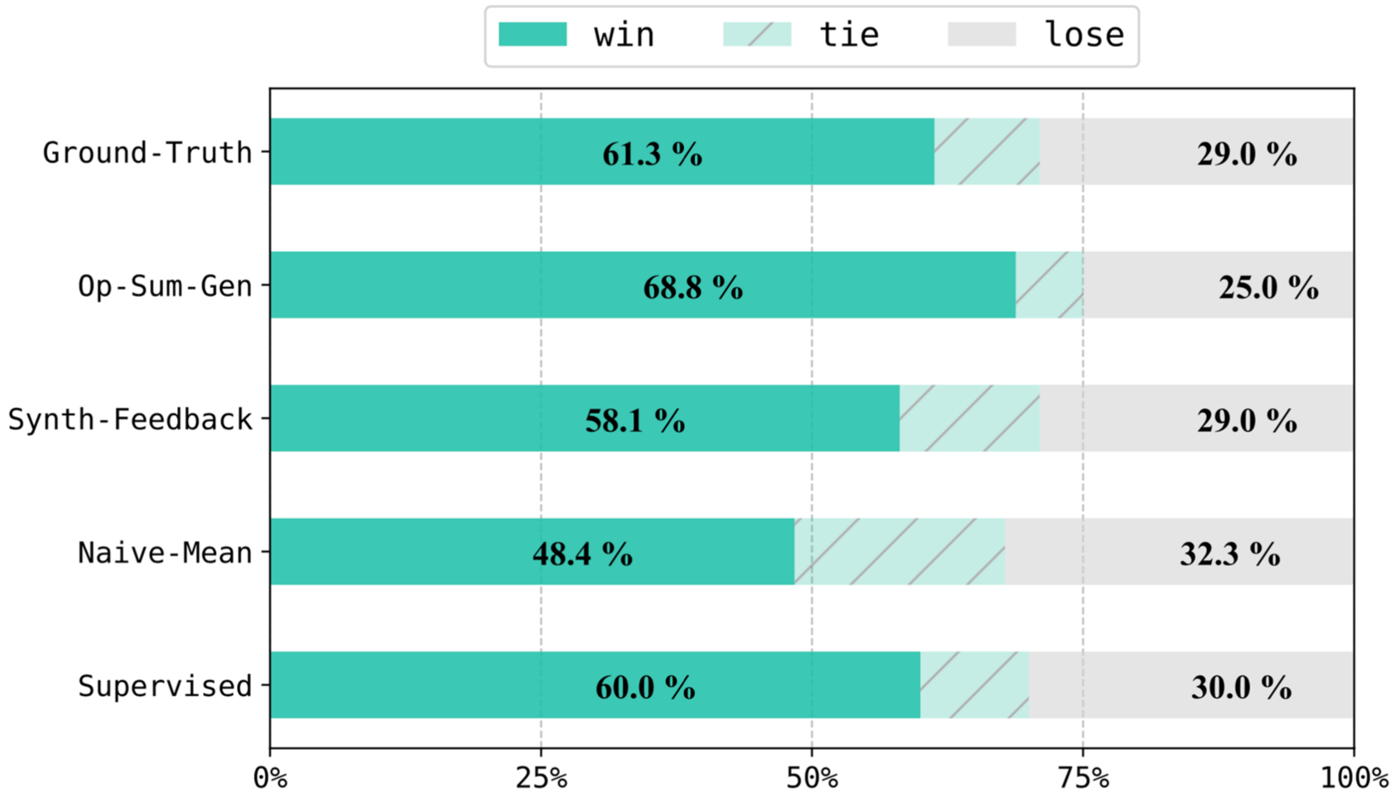
Preference Annotations for only **940** pairs!

> 20x reduction in dataset size!

Set of features:

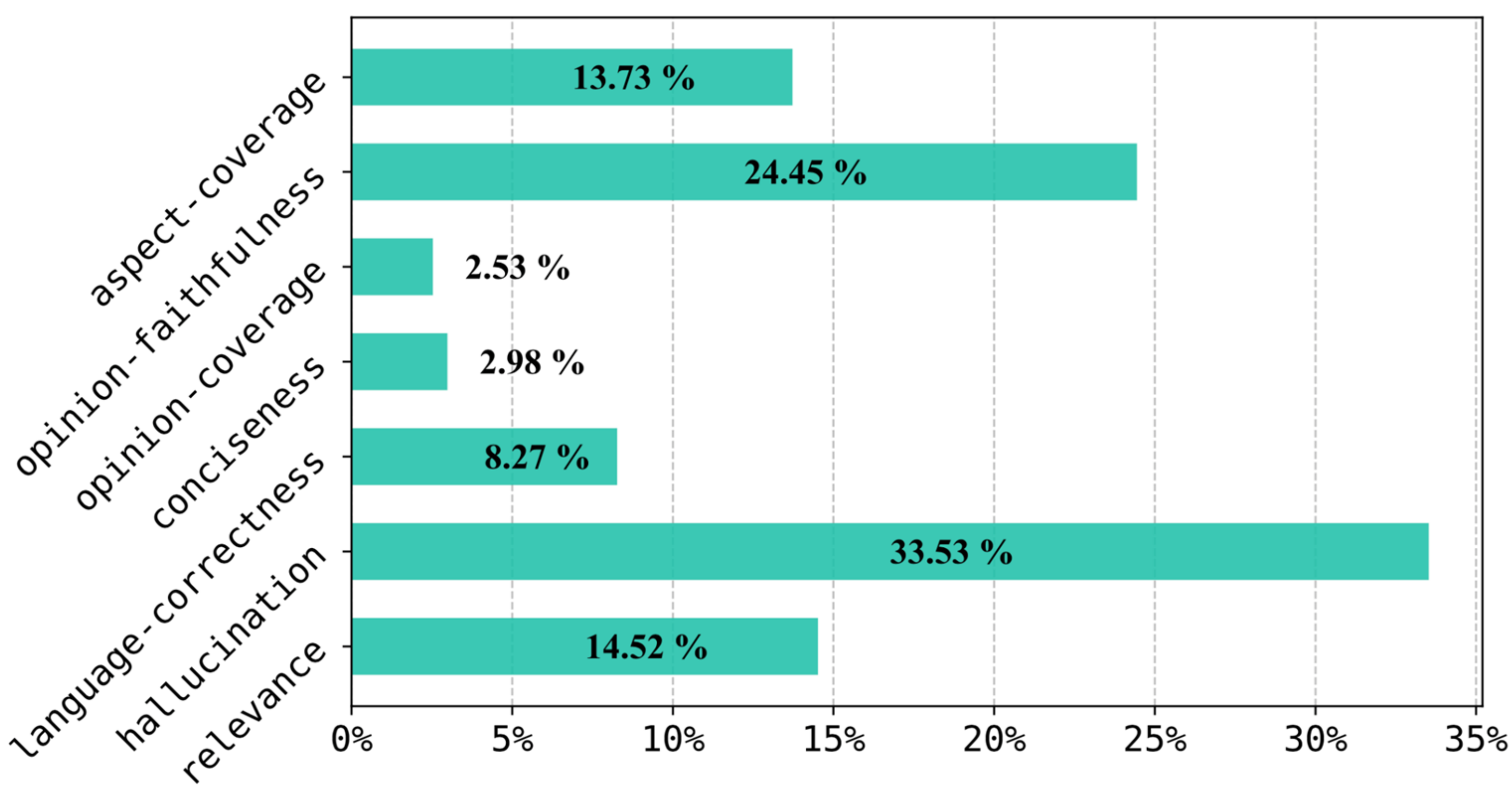
- Aspect Coverage
- Opinion Faithfulness
- Opinion Coverage
- Conciseness
- Relevance
- Hallucination
- Language Correctness

Results



win
tie
lose
Ground-Truth
Op-Sum-Gen
Synth-Feedback
Naive-Mean
Supervised
61.3 %
29.0 %
68.8 %
25.0 %
58.1 %
29.0 %
48.4 %
32.3 %
60.0 %
30.0 %
0%
25%
50%
75%
100%

Analysis


aspect-coverage
opinion-faithfulness
opinion-coverage
conciseness
language-correctness
hallucination
relevance
13.73 %
24.45 %
2.53 %
2.98 %
8.27 %
33.53 %
14.52 %
0%
5%
10%
15%
20%
25%
30%
35%

Conclusion

We show a new way to do RLHF, by leveraging domain knowledge

We show one way to reduce human preference requirement (>20x)

We see that our approach leads to better performance

We see that our approach actually provides alignment too

THANK YOU!!

Pushpak Bhattacharyya

Harshad Khadilkar

Swaroop Nath

Back to Alignment

Derivation of EM based Alignment Expressions

$$V_E = \text{vocalbulary of language } L_1 \text{ (Say English)}$$

$$V_F = \text{vocabulary of language } L_2 \text{ (Say Hindi)}$$

E1 *what is in a name ?*

नाम में क्या है ?

F1 *naam meM kya hai ?*

name in what is ?

E2 *That which we call rose, by any other name will smell as sweet.*

जिसे हम गुलाब कहते हैं, और भी किसी नाम से उसकी कुशबू समान मीठा होगी

F2 *Jise hum gulab kahte hai, aur bhi kisi naam se uski khushbu samaan mitha hogii*

That which we rose say , any other name by its smell as sweet

That which we call rose, by any other name will smell as sweet.

Vocabulary mapping

Vocabulary

V_E	V_F
what , is , in, a , name , that, which, we , call ,rose, by, any, other, will, smell, as, sweet	*naam, meM, kya, hai, jise, ham, gulab, kahte, aur, bhi, kisi, bhi, uski, khushbu, saman, mitha, hogii*

Key Notations

English vocabulary : V_E
French vocabulary : V_F
No. of observations / sentence pairs : S

$$e^1{}_1, e^1{}_2, \dots, e^1{}_{l^1} \Leftrightarrow f^1{}_1, f^1{}_2, \dots, f^1{}_{m^1}$$

$$e^2{}_1, e^2{}_2, \dots, e^2{}_{l^2} \Leftrightarrow f^2{}_1, f^2{}_2, \dots, f^2{}_{m^2}$$

.....

$$e^s{}_1, e^s{}_2, \dots, e^s{}_{l^s} \Leftrightarrow f^s{}_1, f^s{}_2, \dots, f^s{}_{m^s}$$

.....

$$e^S{}_1, e^S{}_2, \dots, e^S{}_{l^S} \Leftrightarrow f^S{}_1, f^S{}_2, \dots, f^S{}_{m^S}$$

No. words on English side in s^{th} sentence : l^s
No. words on French side in s^{th} sentence : m^s
$index_E(e^s{}_p)$ =Index of English word $e^s{}_p$in English vocabulary/dictionary
$index_F(f^s{}_q)$ =Index of French word $f^s{}_q$in French vocabulary/dictionary

Hidden variables and parameters

Hidden Variables (Z) :

Total no. of hidden variables = $\sum_{s=1}^{S} l^s\, m^s$ where each hidden variable is as follows:

$z_{pq}^{s} = 1$, if in s^{th} sentence, p^{th} English word is mapped to q^{th} French word.

$z_{pq}^{s} = 0$, otherwise

Parameters (*Θ*) :

Total no. of parameters = $|\, V_E| \times |V_F|$, where each parameter is as follows:

$P_{i,j}$ = Probability that i^{th} word in English vocabulary is mapped to j^{th} word in French vocabulary

Likelihoods

Data Likelihood *L(D; Θ)* :

$$L(D;\Theta) = \prod_{s=1}^{S}\prod_{p=1}^{l^s}\prod_{q=1}^{m^s}\left(P_{index_E(e_p^s),index_F(f_q^s)}\right)^{z_{pq}^s}$$

Data Log-Likelihood LL(D; Θ) :

$$LL(D;\Theta) = \sum_{s=1}^{S}\sum_{p=1}^{l^s}\sum_{q=1}^{m^s} z_{pq}^s \log\left(P_{index_E(e_p^s),index_F(f_q^s)}\right)$$

Expected value of Data Log-Likelihood E(LL(D; Θ)) :

$$E(LL(D;\Theta)) = \sum_{s=1}^{S}\sum_{p=1}^{l^s}\sum_{q=1}^{m^s} E(z_{pq}^s)\,\log\left(P_{index_E(e_p^s),index_F(f_q^s)}\right)$$

Constraint and Lagrangian

$$\sum_{j=1}^{|V_F|} P_{i,j} = 1 \;, \forall i$$

$$\sum_{s=1}^{S} \sum_{p=1}^{l^s} \sum_{q=1}^{m^s} E(z_{pq}^s) \, log\left(P_{index_E(e_p^s),index_F(f_q^s)}\right) - \sum_{i=1}^{|V_E|} \lambda_i \left(\sum_{j=1}^{|V_F|} P_{i,j} - 1\right)$$

Differentiating wrt P_{ij}

$$\sum_{s=1}^{S}\sum_{p=1}^{l^s}\sum_{q=1}^{m^s}\delta_{index_E(e_p^s),i}\,\delta_{index_F(f_q^s),j}\left(\frac{E(z_{pq}^s)}{P_{i,j}}\right)-\lambda_i=0$$

$$P_{i,j}=\frac{1}{\lambda_i}\sum_{s=1}^{S}\sum_{p=1}^{l^s}\sum_{q=1}^{m^s}\delta_{index_E(e_p^s),i}\,\delta_{index_F(f_q^s),j}\,E(z_{pq}^s)$$

$$\sum_{j=1}^{|V_F|}P_{i,j}=1=\sum_{j=1}^{|V_F|}\frac{1}{\lambda_i}\sum_{s=1}^{S}\sum_{p=1}^{l^s}\sum_{q=1}^{m^s}\delta_{index_E(e_p^s),i}\,\delta_{index_F(f_q^s),j}\,E(z_{pq}^s)$$

Final E and M steps

M-step

$$P_{i,j} = \frac{\sum_{s=1}^{S} \sum_{p=1}^{l^s} \sum_{q=1}^{m^s} \delta_{index_E(e_p^s),i}\, \delta_{index_F(f_q^s),j} E(z_{pq}^s)}{\sum_{j=1}^{|V_F|} \sum_{s=1}^{S} \sum_{p=1}^{l^s} \sum_{q=1}^{m^s} \delta_{index_E(e_p^s),i}\, \delta_{index_F(f_q^s),j} E(z_{pq}^s)}, \forall i, j$$

E-step

$$E(z_{pq}^s) = \frac{P_{index_E(e_p^s),index_F(f_q^s)}}{\sum_{q'=1}^{m^s} P_{index_E(e_p^s),index_F(f_{q'}^s)}}, \forall s, p, q$$

Self Attention

Linguistics of

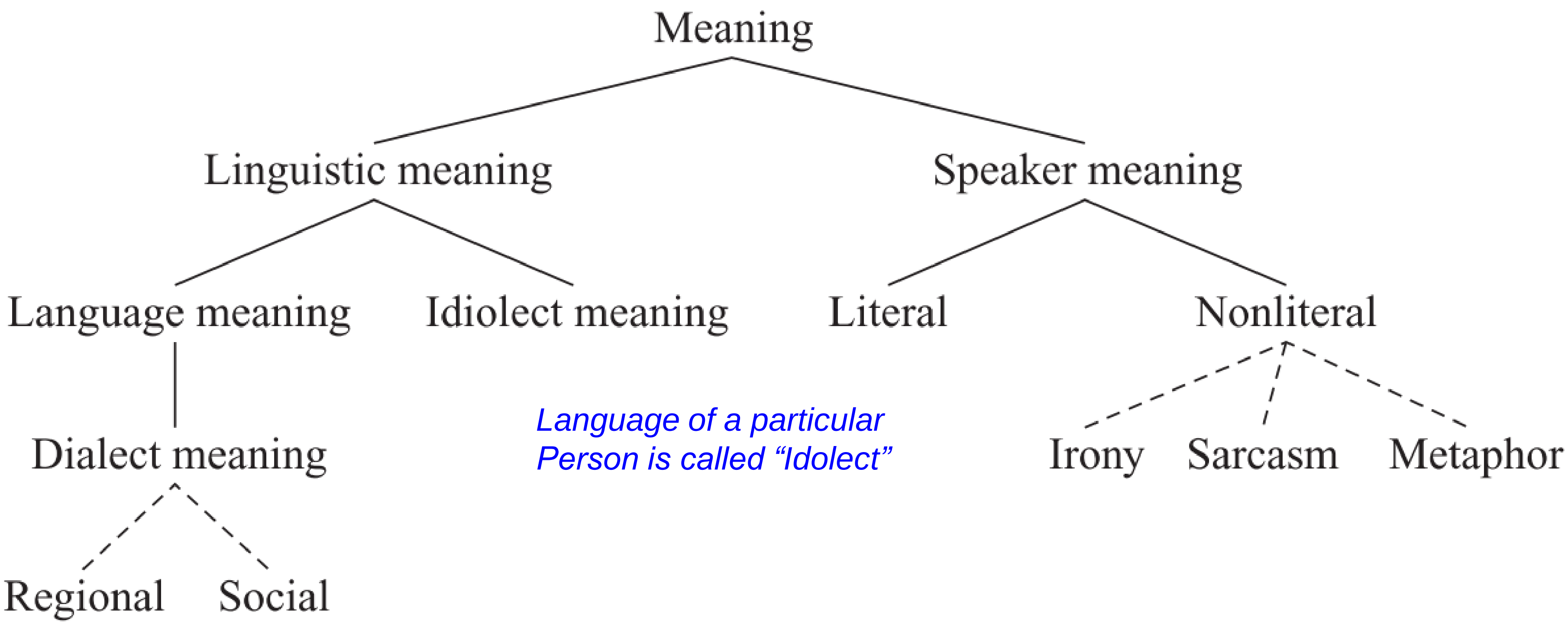
Sound-Structure-Meaning continuum

Sound:
Phonetics, Phonology

Structure:
Morphology, Syntax

Meaning:
Semantic, Pragmatics

The “meaning” ontology



Language of a particular Person is called “Idolect”

From: Akmajian et al. 2010

Meaning of words

Syntagmatic and Paradigmatic Relations

- Syntagmatic and paradigmatic relations
  - Lexico-semantic relations: synonymy, antonymy, hypernymy, mernymy, troponymy etc. **CAT is-a ANIMAL**
  - Coccurence: **CATS MEW**
- Resources to capture semantics:
  - Wordnet: primarily paradigmatic relations
  - ConceptNet: primarily Syntagmatic Relations

Selectional Preferences (Indian Tradition) (1/2)

- “Desire” of some words in the sentence (“aakaangksha”).
  - I saw the boy with long hair.
  - The verb “saw” desires an object here.

- “Appropriateness” of some other words in the sentence to fulfil that desire (“yogyataa”).
  - I saw the boy with long hair.
  - The PP “with long hair” can be appropriately connected only to “boy” and not “saw”.

Selectional Preferences (Indian Tradition) (2/2)

- In case, the ambiguity is still present, “proximity” (“sannidhi”) can determine the meaning.
  - E.g. I saw the boy with a telescope.
  - The PP “with a telescope” can be attached to both “boy” and “saw”, so ambiguity still present. It is then attached to “boy” using the proximity check.

Selectional Preference (Recent Linguistic Theory) (1/2)

- There are words which demand arguments, like, verbs, prepositions, adjectives and sometimes nouns. These arguments are typically nouns.

- Arguments must have the property to fulfil the demand. They must satisfy Selectional preferences.

Selectional Preference (Recent Linguistic Theory) (2/2)

- Example
  - Give (verb)
    - » agent – animate
    - » obj – direct
    - » obj – indirect
  - I gave him the book
  - I gave him the book (yesterday in the school)
    → adjunct

Argument frame and Selectional Preference

- Structure expressing the desire of a word is called the Argument Frame

- Selectional Preference
  - Properties of the “Words that meet the demand”

Verb Argument frame (example)

- Verb: give
- Give
  - Agent: <the giver>: animate
  - direct object: <the given>
  - indirect object: <the givee>: personifiable

- *I_agent gave a book_dobj to John_iobj*