



Étude du transfert d'information dans les conversations naturelles

JTT - Hossam BOUDRAA

Jeudi 07 Décembre 2023

Plan

- 01 Introduction
- 02 Métriques pour les MLMs
- 03 Métriques pour les LLMs
- 04 Conclusion





01

Introduction



Common ground

- Les interlocuteurs développent dynamiquement leur plan de dialogue la construction d'une connaissance partagée (common ground)
- L'espace de travail est une ressource publiquement partagée :
 - ce que A a dit, ce que B a dit, un contexte pertinent
- Les conversations informelles ont rarement un objectif fixe.



© Getty Images

Objectif

- Peut-on obtenir une mesure objective sur la connaissance partagée dans les conversations en français?

Enjeux

- Les modèles de langage cachés (MLMs) pourraient être adaptés à cette nouvelle tâche
- Les larges modèles de langage (LLMs) pour capturer de manière plus complexe et contextuelle les informations échangées

Encadrants



Eliot MAES

Doctorant
LSD



Leonor BECERRA

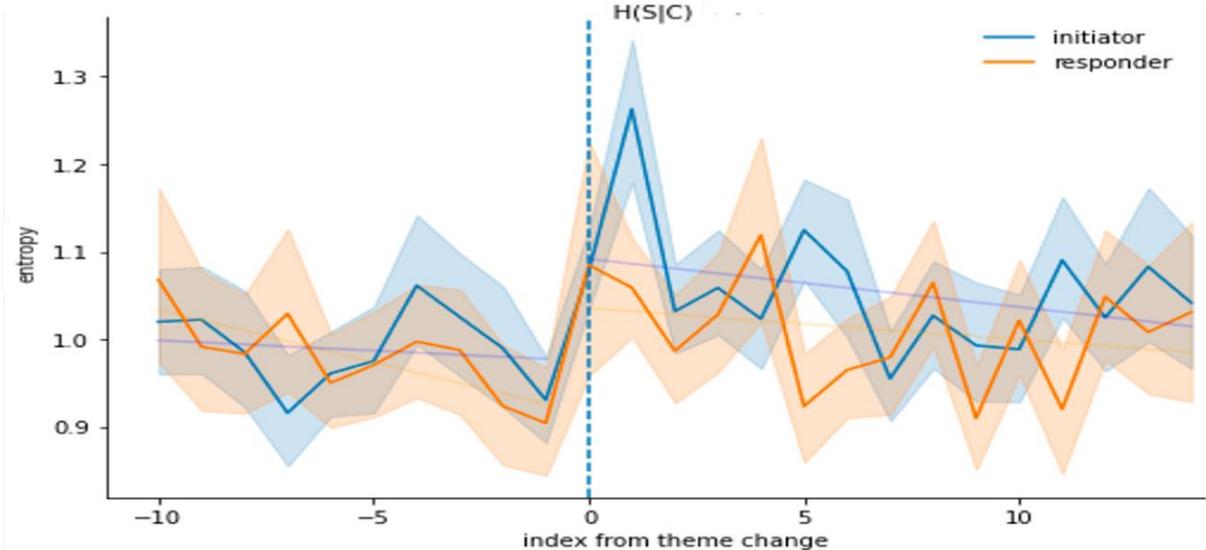
Maîtresse de conférences
Laboratoire d'Informatique et Systèmes



Philippe BLACHE

Directeur de recherche CNRS
Laboratoire Parole et Langage

Shared knowledge in natural conversations: can entropy metrics shed light on information transfers?



Eliot Maës, Philippe Blache, and Leonor Becerra. 2022. (CoNLL)

Dataset

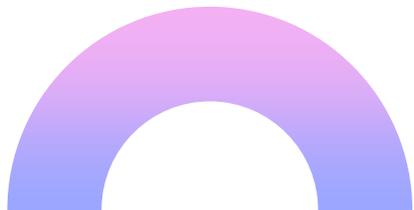
- Paco-Cheese (Amoyal, 2021; LPL, 2018)

2 corpus multimodal contenant des enregistrements (textes, audio et vidéo) de plus de 207 conversations entre participants.

index	theme	speaker	text	H(SIC)	H(S)	MI
120	exams	OR	ça venait de la psycho de l'anthropo enfin de plein de euh domaines	1.18	1.30	0.13
121	exams	AA	ouais ouais	0.31	0.43	0.12
122	exams	OR	je pense c'était juste simplement euh ça	1.04	1.01	-0.04
123	exams	AA	ben ouais mais moi le truc c'est que genre la veille ben du coup je l'avais revu et tout	0.76	0.77	0.01
124	exams	OR	et	0.40	0.62	0.21
125	exams	AA	genre les j'ai vu mes deux résumés je les ai regardés j'ai fait euh	0.95	0.91	-0.03
126	exams	AA	pouah c'est bon ça tombera non non j'ai fait non c'est bon ça tombera pas sur ça	0.76	0.74	-0.01
127	exams	OR	la flemme	1.38	1.03	-0.35
128	exams	OR	ouais	0.37	0.65	0.28
129	exams	AA	genre du coup je les ai lu vite fait en diagonale	1.51	1.55	0.04
130	exams	AA	et	0.51	0.62	0.11
131	exams	AA	après j'ai eu la première question du partiel j'ai fait	0.80	0.85	0.06
132	exams	AA	ah	0.74	0.79	0.05
133	exams	OR	ouais voilà	0.65	0.65	-0.00
134	exams	AA	ah bon	1.41	1.12	-0.29
135	exams	OR	et moi je m'étais même pas rendue compte que c'était là-dedans c'est après cristèle elle m'a dit mais tu vois que c'est tu as exactement ton résumé genre	0.80	0.88	0.08
136	exams	AA	j'aurais dû	0.78	0.75	-0.03
137	exams	OR	alors qu'en plus le résumé quand elle a corrigé il m'a dit très bon résumé	1.11	1.16	0.06
138	exams	AA	ouais moi aussi	0.72	0.76	0.04
139	exams	AA	du coup j'étais un petit peu deg quoi	1.77	1.57	-0.20

02

Métriques pour les MLMs



MLMs : FlauBERT et CamemBERT

- Tokenizer:

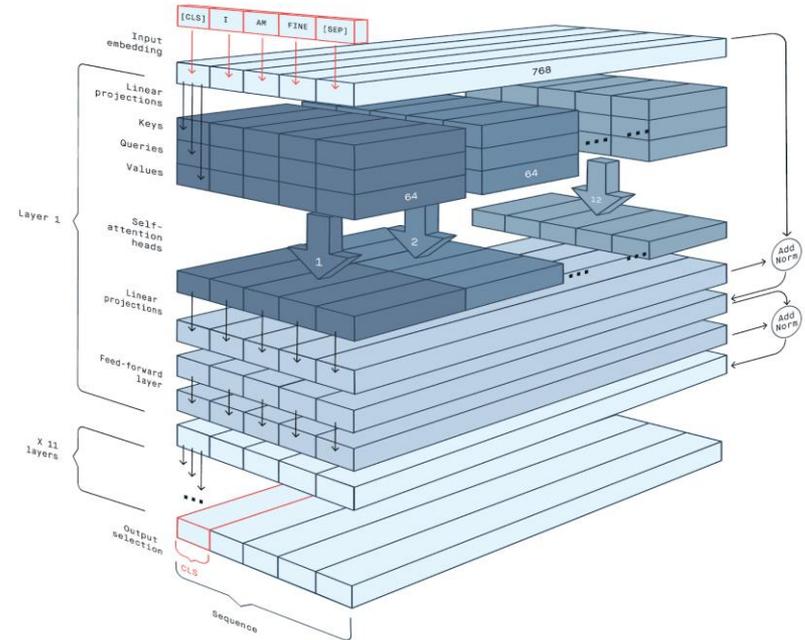
CamemBERT : SentencePiece

FlauBERT : BPE

- Stratégie de masquage:

CamemBERT : Dynamique + Masquage de mots entier

FlauBERT : Dynamique + Masquage de sous-mots



Métriques

- Entropie

```
metric == "WordEntropies":  
metricOutput = -np.sum(word_probs * np.log(word_probs), axis = 1)
```

- Similarité

```
metric == "EmbeddingCosines":  
#if LMOutput.hidden_states is not None:  
if LMOutput[0] is not None:  
    #embeddings = LMOutput.hidden_states[7].detach().numpy()[0]  
    embeddings = LMOutput[0].detach().numpy()[0]  
    metricOutput = [0]  
    for i in range(embeddings.shape[0]-1):  
        metricOutput.append(cosine(embeddings[i], embeddings[i+1]))
```

- Dissimilarité
(Divergence)

```
elif metric == "KLDivergence":  
    metricOutput = [0]  
    for i in range(word_probs.shape[0] - 1):  
        metricOutput.append(-np.sum(word_probs[i + 1] * np.log(word_probs[i])))
```

Entropie

- **CamemBERT-FT**

	theme	Moyenne_Entro
20	classe d'enseignement	6.672535
12	affectation enseignement	6.298767
91	le cours de stats de l'année passée	6.117147
147	les jeux d'ordinateur	6.064678
29	cursus universitaire MA	6.030623
32	date concours et contenu CRPE	5.678138
104	le mode d'examen en social	5.615368
114	le prof de TD de stats	5.590622
117	le programme de social	5.551483
27	cursus actuel CAPES	5.368250
87	le cours Piks	5.316387
133	les barbies customisées	5.288332
187	parcours universitaire des deux participantes	5.268106
174	lieu de rattachement : rapprochement de conjoint	5.238671
149	les lapsus et faux souvenirs	5.226825
177	niveau de connaissance ARTHUS	5.211159
80	laura amie de PC	5.169884
17	but expérience	5.144338
96	le dossier de cognitif	5.130632
155	les partiels d'après	5.126191

- **FlauBERT -FT**

	theme	Moyenne_Entro
35	ergonome	9.564346
57	l'examen de phonologie	9.194622
15	blague AG	9.184373
89	le cours de développement	9.091895
121	le projet professionnel de PO	9.027339
39	explication de la blague de AD	8.779411
172	lieu de mobilisation armée	8.512864
103	le master de psycho	8.444117
93	le cursus actuel de BE et le projet pro	8.391385
17	but expérience	8.265492
24	comprendre les rêves	8.141446
31	dangerosité de l'armée	8.056571
90	le cours de projet professionnel	7.976916
72	la présentation en anglais de CM	7.906467
13	auto selection	7.874297
4	Facebook	7.721865
23	commentaires sur les textes	7.720501
141	les enjeux des familles nombreuses	7.713831
134	les camarades de LS	7.709517
41	explicitation des consignes	7.693721

Similarité

- CamemBERT-FT

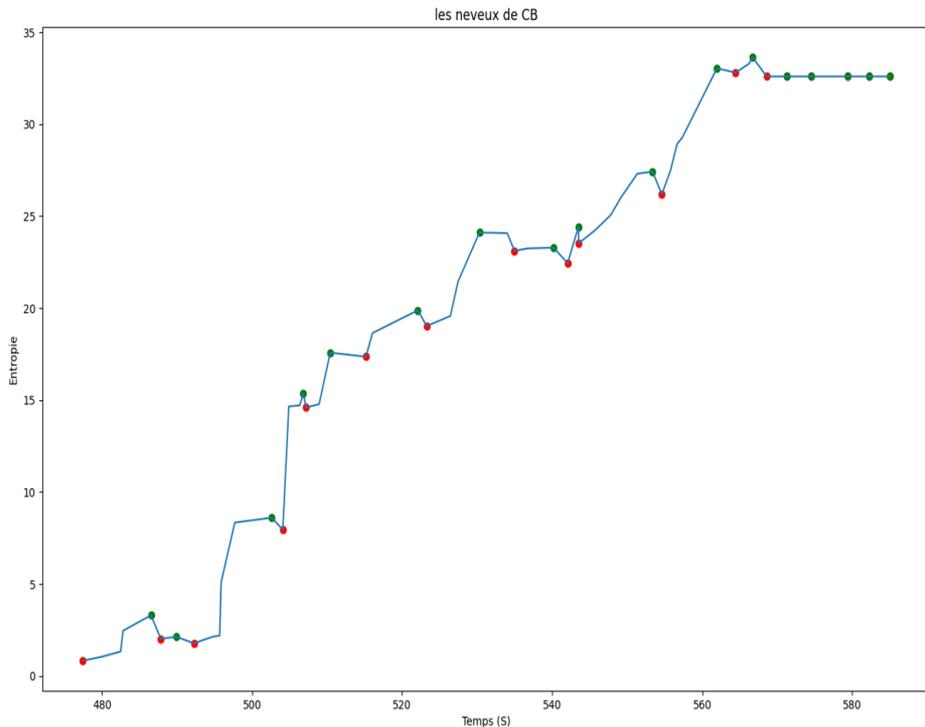
	theme	Moyenne_Sim
20	classe d'enseignement	3.148688
29	cursus universitaire MA	2.963457
12	affectation enseignement	2.779732
91	le cours de stats de l'année passée	2.729442
147	les jeux d'ordinateur	2.705207
187	parcours universitaire des deux participantes	2.558622
177	niveau de connaissance ARTHUS	2.516539
104	le mode d'examen en social	2.477086
149	les lapsus et faux souvenirs	2.461262
114	le prof de TD de stats	2.404795
155	les partiels d'après	2.316391
117	le programme de social	2.279978
133	les barbies customisées	2.272687
27	cursus actuel CAPES	2.262324
116	le programme de neurologie	2.236777
154	les partiels bloqués	2.187414
17	but expérience	2.169494
100	le langage et le corps : le théâtre	2.148897
87	le cours Piks	2.143533
131	les amis d'enfance	2.108596
89	le cours de développement	2.092238

- FlauBERT-FT

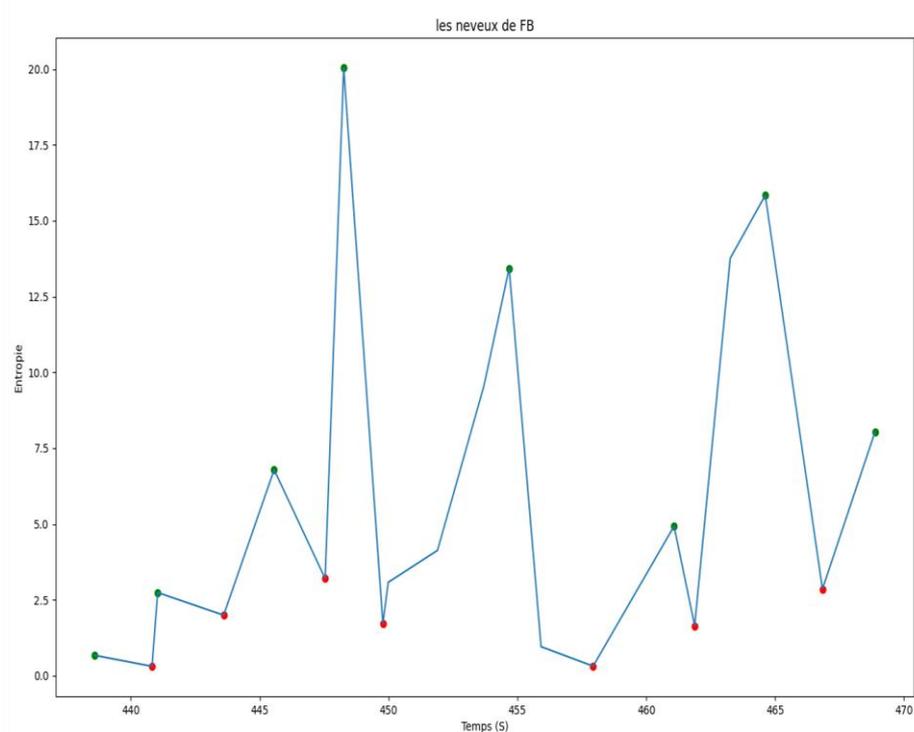
	theme	Moyenne_Sim
20	classe d'enseignement	5.048897
29	cursus universitaire MA	4.953694
12	affectation enseignement	4.794067
91	le cours de stats de l'année passée	4.778155
114	le prof de TD de stats	4.201252
187	parcours universitaire des deux participantes	4.063995
177	niveau de connaissance ARTHUS	3.940607
116	le programme de neurologie	3.756054
104	le mode d'examen en social	3.691622
131	les amis d'enfance	3.686286
149	les lapsus et faux souvenirs	3.584847
133	les barbies customisées	3.579511
147	les jeux d'ordinateur	3.548315
174	lieu de rattachement : rapprochement de conjoint	3.472222
154	les partiels bloqués	3.426802
27	cursus actuel CAPES	3.424944
87	le cours Piks	3.380014
180	organisation avant vacances	3.271748
171	lieu d'origine de LS	3.254136
6	Londres	3.173017
117	le programme de social	3.161479

Moments de pic /thème

- **CamemBERT -FT**

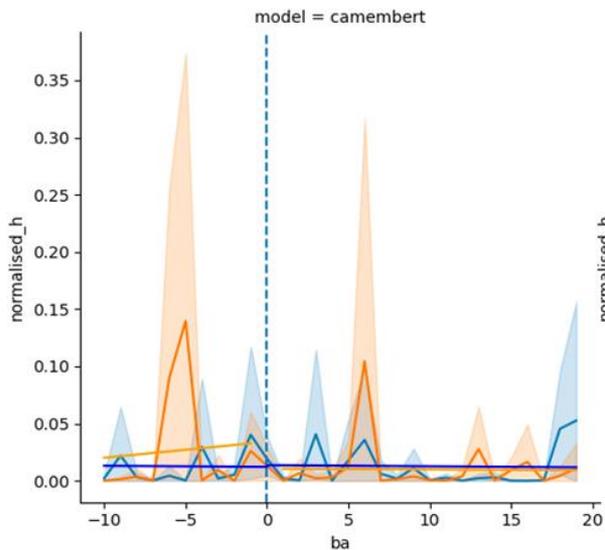


- **FlauBERT -FT**



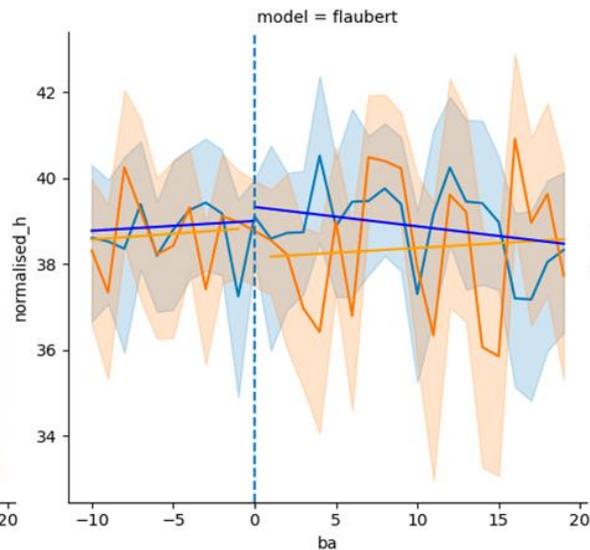
Comparatif

- **CamemBERT -FT**



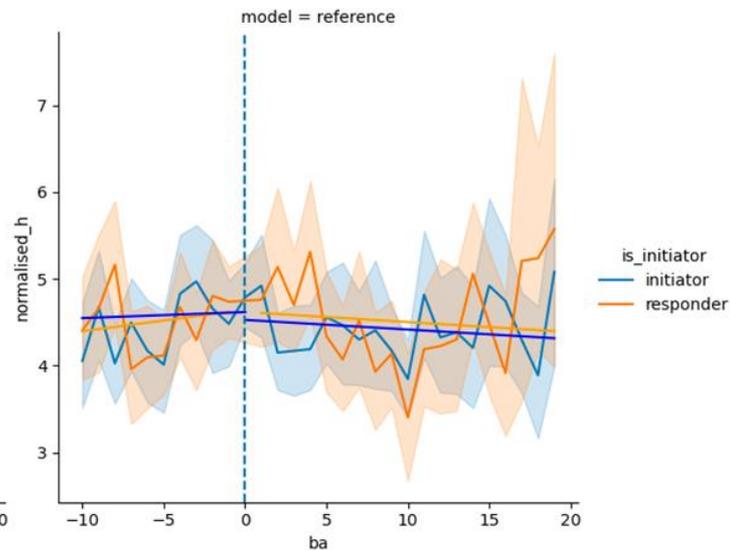
CamemBERT - GPT2 (IC* = 60%)

- **FlauBERT -FT**



FlauBERT - GPT2 (IC* = 90%)

- **GPT2 -FT**



* : Intervalle de confiance

03

Métriques pour les LLMs



Métriques LLMs:

❖ Métriques pour les résumés automatiques:

- ROUGE-1, ROUGE-2, ROUGE-L et ROUGE-S

❖ Métriques pour les transferts d'informations :

- $\text{ann}_c, \text{ann}_s$
- $\text{Semb}_{S \rightarrow \text{ipu}}$
- $\text{SUMMAC}_{ZS}, \text{SUMMAC}_{\text{CONV}}$

Les résumés automatiques

BARTv1	T5	GPT 2	BARTv2	ChatGPT
<p>Demande de réunion avec un client très important. L'agent répond qu'il y a une campagne importante et le temps perdu après la semaine prochaine.</p>	<p>Il y a une campagne importante la semaine prochaine qui me tient occupé, sinon le reste va bien dans ma vie.</p>	<p>Harry : Oui, tu as raison, ça fait longtemps qu'on ne s'est pas vus. Je viens de terminer une réunion avec un client très important et maintenant j'ai enfin du temps libre.</p>	<p>ève de réunion avec un client très important et maintenant du temps libre. Faisons un plan et rattrapons le temps perdu après la semaine prochaine. Au revoir, prends soin de toi.</p>	<p>Rann et Harry se rencontrent après une longue période sans se voir. Rann est occupé avec une campagne importante, tandis que Harry vient de terminer une réunion avec un client. Ils décident de planifier une rencontre pour rattraper le temps perdu après la semaine prochaine. Ils se disent au revoir et se souhaitent mutuellement de prendre soin d'eux.</p>

Approche extractive

Conversation	Résumé
<p>Rann : Hé Harry, comment ça va ? Ça fait longtemps qu'on ne s'est pas vus !</p> <p>Harry : Hé ! Quelle surprise !</p> <p>Harry : Oui, tu as raison, ça fait longtemps qu'on ne s'est pas vus. Comment ça va ?</p> <p>Rann : Il y a une campagne importante la semaine prochaine qui me tient occupé, sinon le reste va bien dans ma vie.</p> <p>Rann : Et toi ?</p> <p>Harry : Oh ! Je viens de terminer une réunion avec un client très important et maintenant j'ai enfin du temps libre. Je suis soulagé d'en avoir fini avec ça.</p> <p>Rann : C'est bien pour toi alors. Hé ! Faisons un plan et rattrapons le temps perdu après la semaine prochaine.</p> <p>Rann : Qu'en dis-tu ?</p> <p>Harry : Bien sûr, pourquoi pas ? Appelle-moi quand tu auras terminé ton projet.</p> <p>Rann : D'accord, alors.</p> <p>Rann : Au revoir, prends soin de toi.</p> <p>Harry : Au revoir, mon pote.</p>	<p>Rann et Harry se rencontrent après une longue période sans se voir. Rann est occupé avec une campagne importante, tandis que Harry vient de terminer une réunion avec un client. Ils décident de planifier une rencontre pour rattraper le temps perdu après la semaine prochaine. Ils se disent au revoir et se souhaitent mutuellement de prendre soin d'eux.</p> <hr/> <p>Harry : Oui, tu as raison, ça fait longtemps qu'on ne s'est pas vus. Je viens de terminer une réunion avec un client très important et maintenant j'ai enfin du temps libre.</p>

ChatGPT

GPT2

Approche abstraactive

Conversation	Résumé
<p>Rann : Hé Harry, comment ça va ? Ça fait longtemps qu'on ne s'est pas vus !</p> <p>Harry : Hé ! Quelle surprise !</p> <p>Harry : Oui, tu as raison, ça fait longtemps qu'on ne s'est pas vus. Comment ça va ?</p> <p>Rann : Il y a une campagne importante la semaine prochaine qui me tient occupé, sinon le reste va bien dans ma vie.</p> <p>Rann : Et toi ?</p> <p>Harry : Oh ! Je viens de terminer une réunion avec un client très important et maintenant j'ai enfin du temps libre. Je suis soulagé d'en avoir fini avec ça.</p> <p>Rann : C'est bien pour toi alors. Hé ! Faisons un plan et rattrapons le temps perdu après la semaine prochaine.</p> <p>Rann : Qu'en dis-tu ?</p> <p>Harry : Bien sûr, pourquoi pas ? Appelle-moi quand tu auras terminé ton projet.</p> <p>Rann : D'accord, alors.</p> <p>Rann : Au revoir, prends soin de toi.</p> <p>Harry : Au revoir, mon pote.</p>	<p>Rann et Harry se rencontrent après une longue période sans se voir. Rann est occupé avec une campagne importante, tandis que Harry vient de terminer une réunion avec un client. Ils décident de planifier une rencontre pour rattraper le temps perdu après la semaine prochaine. Ils se disent au revoir et se souhaitent mutuellement de prendre soin d'eux.</p> <hr/> <p>Rann est très content de revoir Harry. Il est occupé actuellement par une importante campagne. Harry vient de finir une réunion avec un client, il est satisfait.</p> <p>Rann souhaite rattraper le temps perdu après la semaine prochaine.</p> <p>Rann a bien reçu la réaction de Harry.</p>

ChatGPT

LLaMA

Comparatif :

BART	ChatGPT
Demande de réunion avec un client très important et maintenant du temps libre. Faisons un plan et rattrapons le temps perdu après la semaine prochaine. Au revoir, prends soin de toi.	Rann et Harry se rencontrent après une longue période sans se voir. Rann est occupé avec une campagne importante, tandis que Harry vient de terminer une réunion avec un client. Ils décident de planifier une rencontre pour rattraper le temps perdu après la semaine prochaine. Ils se disent au revoir et se souhaitent mutuellement de prendre soin d'eux.

Rouge1	Rouge2	RougeL	Rougesum
42.86	25	43.86	34.70

Comparatif :

LLaMA 13B	ChatGPT
<p>Rann est très content de revoir Harry. Il est occupé actuellement par une importante campagne.</p> <p>Harry vient de finir une réunion avec un client, il est satisfait.</p> <p>Rann souhaite rattraper le temps perdu après la semaine prochaine.</p> <p>Rann a bien reçu la réaction de Harry.</p>	<p>Rann et Harry se rencontrent après une longue période sans se voir. Rann est occupé avec une campagne importante, tandis que Harry vient de terminer une réunion avec un client. Ils décident de planifier une rencontre pour rattraper le temps perdu après la semaine prochaine. Ils se disent au revoir et se souhaitent mutuellement de prendre soin d'eux.</p>

Rouge1	Rouge2	RougeL	Rougesum
52.63	30.36	43.86	43.86

Évaluation des résumés automatiques

Sur 207 conversations/thèmes :

Source	ChatGPT	Vigogne 13b
Rouge-1	0.25 ± 0.08	0.13 ± 0.08
Rouge-2	0.08 ± 0.04	0.01 ± 0.02
Rouge-L	0.17 ± 0.05	0.09 ± 0.05

temp: .1 / .3 / .5 / .7 / .9

max tokens : 1024 / 2048

Annotation manuelle des IPU

Niveau	Descriptif
1	IPU informative
2	IPU informative secondaire
3	IPU non informative



Niveau	Descriptif
ann _s	IPU informative
ann _c	IPU non informative

Semb_{S→ipu}

1. Étiquetage morpho-syntaxique (POS)
2. Extraction des enchâssements (embeddings) contextualisés
3. Normalisation
4. Calcul de la similarité entre chaque IPU et phrase de résumé
5. $\text{Max} | \text{cosine_sim}\{S_{ik}, \text{IPU}_{jk}\} \mid i, j \in N^*M, k \in \{1, 207\}$

Semb_{S→ipu}

Thème : **Lieu d'habitation**

Phrases

S1 :La conversation se déroule entre plusieurs personnes, chacune ayant une contribution différente.

S2:La personne A (AA) discute avec la personne B (OR) de son logement et de ses options financières.

...

IPUs

AA:Mh

MH: pou plutôt sympa euh.

OR: tu imagines presque 500 balles pour j'ai un studio genre c'est vraiment et en plus on est deux dedans.

OR: tu payes 485 euros pour 24 mètres carré

0.0060

0,98

0,95

SUMMAC_{ZS}, SUMMAC_{CONV}

Sentence-Level NLI

$$P(Y = entail | D_i, S_j)$$

Document

Scientists are studying Mars to learn about the Red Planet and find landing sites for future missions. **D1**

One possible site, known as Arcadia Planitia, is covered in strange sinuous features. **D2**

The shapes could be signs that the area is actually made of glaciers, which are large masses of slow-moving ice. **D3**

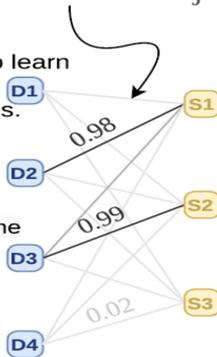
Arcadia Planitia is in Mars' northern lowlands. **D4**

Summary

There are strange shape patterns on Arcadia Planitia. **S1** ✓

The shapes could indicate the area might be made of glaciers. **S2** ✓

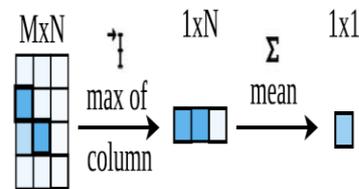
This makes Arcadia Planitia ideal for future missions. **S3** ✗



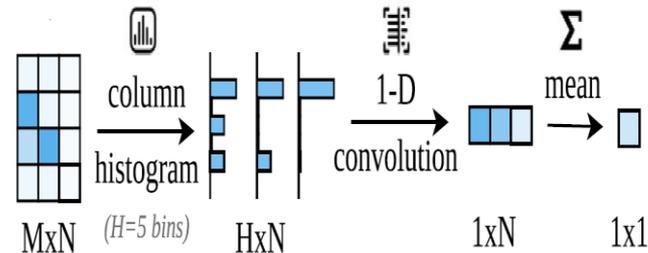
Document-Level NLI

$$P(Y = entail | document, summary) = 0.91$$

$$X_{pair} = \begin{bmatrix} 0.02 & 0.02 & 0.04 \\ 0.98 & 0.00 & 0.00 \\ 0.43 & 0.99 & 0.00 \\ 0.00 & 0.00 & 0.01 \end{bmatrix}$$



SUMMAC_{ZS}



SUMMAC_{CONV}

- **Localisation des informations importantes**
- **Résultats finals**

JTT Eliot

04 Conclusion



Conclusion

01

Métriques pour MLMs

- FlauBERT et CamemBERT obtiennent des résultats différentes pour l'entropie
- Corrélation de l'entropie entre Flaubert et le modèle préexistant (IC:90%)

02

Métriques pour LLMs

- Les résultats de Vigogne 13b n'ont pas atteint celles de ChatGPT
- L'efficacité de la métrique $\text{Semb}_{S \rightarrow \text{ipu}}$ reste encore à analyser avec d'autres LLMs

Références

- Eliot Maës, Philippe Blache, and Leonor Becerra Bonache. 2022. Shared knowledge in natural conversations: can entropy metrics shed light on information transfers? In 26th Conference on Computational Natural Language Learning (CoNLL), pages 213–227.
- Martin Pickering and Simon Garrod. 2021. Understanding Dialogue. Cambridge University Press.
- Martin J. Pickering and Simon Garrod. 2004. Toward a mechanistic psychology of dialogue. Behavioral and Brain Sciences, 27(2):169–190.
- Philippe Laban, Tobias Schnabel, Paul N. Bennett, Marti A. Hearst; SummaC: Re-Visiting NLI-based Models for Inconsistency Detection in Summarization. Transactions of the Association for Computational Linguistics 2022; 10 163–177
- Béchet, Frédéric, Benjamin Maza, Nicolas Bigouroux, Thierry Bazillon, Marc El-Bèze, Renato De Mori and Eric Arbillot. “DECODA: a call-centre human-human spoken conversation corpus.” International Conference on Language Resources and Evaluation (2012)

Merçi!

Hossam BOUDRAA

hossam.boudraa@lis-lab.fr

