Sequence-to-Sequence Language Models for Character and Emotion Detection in Dream Narratives

Gustave Cortal

Laboratoire Méthodes Formelles

école normale supérieure paris–saclay

université PARIS-SACLAY

What is the function of dreaming?

Dreams might function as:

- expressions of desires repressed during the waking state [6]
- emotional conflict resolution [1, 9]
- memory consolidation, selectively forgetting irrelevant information to facilitate learning [3, 2]
- simulators that train individuals to better react to new situations [8]

$\rightarrow$ We still don't have a definite answer due to a lack of scientific evidence

Quantitative dream analysis

Scientific evidence for the **continuity hypothesis**: dreams prolong events experienced during waking [7]

Quantitative dream analysis has studied this hypothesis, emphasizing the quantitative aspects by examining recurring patterns and associations in dream narratives [11, 4]

$\rightarrow$ We need a lot of dream narratives and an annotation scheme

Hall and Van de Castle annotation scheme

The **Hall and Van de Castle** (HVdC) scheme identifies characters, emotions, interactions, and objects present in the narratives [10]

Applying this scheme to narratives is time-consuming due to its complexity and the need to train annotators for *manual* dream analysis

Although many available narratives exist, only a small subset has been annotated according to this scheme [5]

→ Automating the coding of narratives is an important challenge, as it would accelerate dream research by making thousands of annotated narratives available

DreamBank, a corpus of dream narratives

series	information	years	nb
ed	adult man	1980-2002	143
bea1	teenager girl	2003-2005	136
b-baseline	adult woman	1960-1997	234
emma	adult woman	1949-1997	285
norms-m	adult men	1940s-1950s	485
norms-f	adult women	1940s-1950s	483

Figure: Series of dreamers with the number of annotated narratives.

DreamBank consists of 27,000+ dream narratives mostly in English

We use a subset of 1,766 narratives annotated according to the HVdC scheme for training and evaluation

Our approach

In DreamBank, narratives are annotated according to the HVdC scheme, which codes characters and their emotions with symbols

We use the HVdC annotation guidelines to map each symbol to its corresponding linguistic label

We convert these symbols into the natural language to better exploit the semantics of their references using language models

We jointly identify characters and their emotions using a sequence-to-sequence approach

Our approach

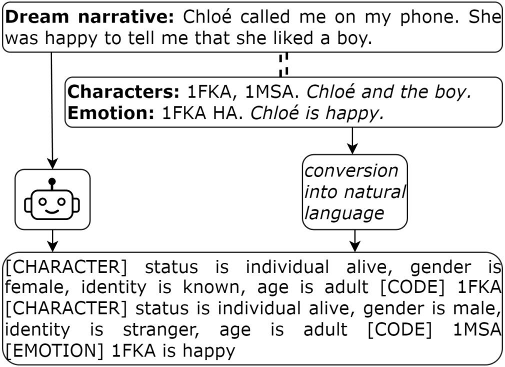


Figure: Sequence-to-sequence approach for automating the coding of dream narratives. Codes describing characters and their emotions are converted into natural language to produce the training data. From a narrative, a language model generates the natural language description of characters and their emotions.

Example

Narrative: It was my birthday and I was having a party but in a place I've never been before. It was in a forest type area. All I remember is that at the same time I had two boyfriends. Only one was at my party, though he had just broken up with my best friend so I kinda felt uncomfortable being with him. We had got in an argument so he left. I don't quite remember how but we did make up but I don't remember when or why even got in the argument. I woke up when I heard the telephone ringing.
Coding: 2MSC, 1MSC, 1FSC D, AP

Figure: A narrative and its coding.

Conversion for characters

- Status: individual alive (1), group alive (2), dead individual (3), dead group (4), imaginary individual (5), imaginary group (6), original form (7), changed form (8)
- Gender: male (M), female (F), joint (J), indefinite (I)
- Identity: known (K), prominent (P), occupational (O), ethnic (E), stranger (S)
- Age: adult (A), child (C)

Annotation guidelines for characters and emotions are available at `https://dreams.ucsc.edu/Coding/`

Conversion for characters

The coding of "[...] I had two boyfriends. Only one was at my party, though he had just broken up with my best friend so I kinda felt uncomfortable with him [...]" is "2MSC, 1MSC, 1FSC"

Using the annotation guideline:

"2MSC" is converted to "status is group alive, gender is male, identity is stranger, age is child"

"1MSC" is converted to "status is individual alive, gender is male, identity is stranger, age is child"

"1FSC" is converted to "status is individual alive, gender is female, identity is stranger, age is child"

Sequence-to-sequence approach for character prediction

Based on the narrative, we need to generate:

[CHARACTER] status is group alive, gender is male, identity is stranger, age is child [CHARACTER] status is individual alive, gender is male, identity is stranger, age is child [CHARACTER] status is individual alive, gender is female, identity is stranger, age is child

During the evaluation, we convert our generated sentence into a sequence of codes:

2MSC, 1MSC, 1FSC

Conversion for emotion

The HVdC scheme considers five emotional states: anger (AN), apprehension (AP), sadness (SD), confusion (CO), and happiness (HA)

The coding of "[...] I kinda felt uncomfortable [...]" is "D, AP"

Using the annotation guideline:

"D, AP" is converted to "dreamer has apprehension"

Sequence-to-sequence approach for emotion prediction

Based on the narrative, we need to generate:

dreamer has apprehension

During the evaluation, we convert our generated sentence into a sequence of codes:

D, AP

Sequence-to-sequence approach for character and emotion prediction

Based on the narrative, we need to generate:

[CHARACTER] status is group alive, gender is male, identity is stranger, age is child [CHARACTER] status is individual alive, gender is male, identity is stranger, age is child [CHARACTER] status is individual alive, gender is female, identity is stranger, age is child [EMOTION] dreamer has apprehension

During the evaluation, we convert our generated sentence into a sequence of codes:

2MSC, 1MSC, 1FSC for character prediction

D, AP for emotion prediction

Experiments

To construct our baseline, we finetune LaMini-Flan-T5 (248M encoder-decoder model) on our annotated narratives

To get insight into prediction performance, we investigate several phenomena, such as the effect of the language model size, character prediction order, the manner of converting codes into natural language, the consideration of proper names, and the semantics of character references

To compare our approach, we perform in-context learning with StableBeluga, a 7B decoder-only model

Results

model	status	gender	identity	age	character	emotion
BASELINE	82.87	78.02	76.17	86.21	64.74	75.13
$\text{NO}_{\text{SEMANTICS}}$	71.37	56.54*	61.0	90.51	41.79*	75.79
NO_{NAMES}	80.66*	74.32**	74.2	83.95*	60.93**	73.04*
$\text{SIZE}_{\text{SMALL}}$	78.35**	72.13**	70.25**	81.66**	56.79**	70.15**
$\text{SIZE}_{\text{LARGE}}$	84.51*	80.3**	78.63**	87.29	67.63**	74.71
$\text{FIRST}_{\text{GROUP}}$	82.33	77.71	74.86	85.61	63.71	71.94
$\text{FIRST}_{\text{INDIVIDUAL}}$	80.59**	76.14	74.22*	83.87**	62.67	67.32
$\text{FIRST}_{\text{EMOTION}}$	83.92	78.74	77.06	87.63	64.97	72.03
$\text{CONVERSION}_{\text{COMMA}}$	84.02**	79.84**	77.67**	87.08*	66.69**	73.68
$\text{CONVERSION}_{\text{MARKER}}$	82.39	78.45	76.53	86.09	65.44	74.36
STABLEBELUGA1	43.95**	39.76**	31.25**	56.16**	15.65**	-
STABLEBELUGA3	52.44**	46.49**	38.46**	63.88**	21.06**	-
STABLEBELUGA5	55.89**	46.29**	42.61**	63.73**	24.86**	-
CROSS-VALIDATION	86.28	81.9	79.51	89.52	68.64	76.18

**: $p < 0.05$
*: $p < 0.1$

How effective is our baseline?

model	status	gender	identity	age	character	emotion
BASELINE	82.87	78.02	76.17	86.21	64.74	75.13

→ Our supervised models can effectively address the complex task of predicting characters and their associated emotions through our approach

How do series-specific factors impact performance?

model	status	gender	identity	age	character	emotion
BASELINE	82.87	78.02	76.17	86.21	64.74	75.13
CROSS-VALIDATION	86.28	81.9	79.51	89.52	68.64	76.18

→ The model relies on the specificities of the training set series to predict new dreams

How does imposing an order in character prediction impact performance?

model	status	gender	identity	age	character	emotion
BASELINE	82.87	78.02	76.17	86.21	64.74	75.13
FIRST$_{\text{GROUP}}$	82.33	77.71	74.86	85.61	63.71	71.94
FIRST$_{\text{INDIVIDUAL}}$	80.59**	76.14	74.22*	83.87**	62.67	67.32

FIRST$_{\text{INDIVIDUAL}}$ predicts individuals before groups, and FIRST$_{\text{GROUP}}$ predicts groups before individuals

→ We can't conclude that a specific order is better

How does the sequence of predicting emotions and characters impact performance?

model	status	gender	identity	age	character	emotion
BASELINE	82.87	78.02	76.17	86.21	64.74	75.13
FIRST$_{\text{EMOTION}}$	83.92	78.74	77.06	87.63	64.97	72.03

FIRST$_{\text{EMOTION}}$ studies the effect of predicting emotions before predicting characters, which involves reversing the two steps in a generation

→ We can't conclude that emotion prediction relies on character prediction

How does using proper names impact performance?

model	status	gender	identity	age	character	emotion
BASELINE	82.87	78.02	76.17	86.21	64.74	75.13
NO_{NAMES}	80.66*	74.32**	74.2	83.95*	60.93**	73.04*

NO_{NAMES}: We apply a named entity recognition model to detect proper names in narratives, which are replaced by the specific token "[PER]"

→ The model relies on proper names to predict characters and their emotions

How do the model sizes impact performance?

model	status	gender	identity	age	character	emotion
BASELINE	82.87	78.02	76.17	86.21	64.74	75.13
$\text{SIZE}_{\text{SMALL}}$	78.35**	72.13**	70.25**	81.66**	56.79**	70.15**
$\text{SIZE}_{\text{LARGE}}$	84.51*	80.3**	78.63**	87.29	67.63**	74.71

$\text{SIZE}_{\text{SMALL}}$ and $\text{SIZE}_{\text{LARGE}}$ contain 77 and 783 million parameters respectively

$\rightarrow$ Scaling up the models is an interesting direction for improving character prediction

How does converting codes into natural language impact performance?

model	status	gender	identity	age	character	emotion
BASELINE	82.87	78.02	76.17	86.21	64.74	75.13
$\text{CONVERSION}_{\text{COMMA}}$	84.02**	79.84**	77.67**	87.08*	66.69**	73.68
$\text{CONVERSION}_{\text{MARKER}}$	82.39	78.45	76.53	86.09	65.44	74.36

BASELINE separates subclasses with linguistic markers (*e.g.*, "status is individual alive, gender is female, identity is known, age is adult")

$\text{CONVERSION}_{\text{COMMA}}$ separates subclasses with commas (*e.g.*, "individual alive, female, known, adult")

$\text{CONVERSION}_{\text{MARKER}}$ separates subclasses with specific markers (*e.g.*, "[STATUS] individual alive [GENDER] female [IDENTITY] known [AGE] adult")

→ It is unnecessary to introduce linguistic markers to link the subclasses

How does StableBeluga perform compared to our supervised models?

model	status	gender	identity	age	character	emotion
baseline	82.87	78.02	76.17	86.21	64.74	75.13
StableBeluga_1	43.95**	39.76**	31.25**	56.16**	15.65**	-
StableBeluga_3	52.44**	46.49**	38.46**	63.88**	21.06**	-
StableBeluga_5	55.89**	46.29**	42.61**	63.73**	24.86**	-

The number of in-context examples increases performance

→ Our supervised models perform better while having 28 times fewer parameters

How does considering the semantic aspects of character references impact performance?

model	status	gender	identity	age	character	emotion
BASELINE	82.87	78.02	76.17	86.21	64.74	75.13
NO$_{\text{SEMANTICS}}$	71.37	56.54*	61.0	90.51	41.79*	75.79

NO$_{\text{SEMANTICS}}$ does not consider the semantics of character references, as it directly predicts the symbols. Revisiting the example in Figure 2, the target text will be "[CHARACTER] 1FKA [CHARACTER] 1MSA [EMOTION] 1MSA is happy"

→ Converting the character codes into natural language allows the language model to leverage their semantics

Contributions

- The joint prediction of characters and their emotions in dream narratives using a sequence-to-sequence language model
- The examination of various phenomena, allowing insight into prediction performance, such as the effect of language model size, the order of character prediction, the conversion of codes into natural language, and the consideration of proper names and character traits
- The comparison of our approach with a large language model using in-context learning. Our supervised models perform better while having 28 times fewer parameters
- The release of our model[1] and the English part of the DreamBank corpus, including 27,952 annotated dream narratives[2]

[1] `https://huggingface.co/gustavecortal/dream-t5`
[2] `https://huggingface.co/datasets/gustavecortal/DreamBank-annotated`

Perspectives

- Include non-annotated dreams, perhaps using unsupervised learning on the narratives
- Consider all available annotations in the HVdC scheme, such as interactions between characters, objects, and settings
- Design a better prompt for dream analysis, perhaps using a question-answering approach

Appendix: distribution of emotional states

Out of 1,766 narratives, 885 have no emotional content. On average, narratives with emotional content have 1.6 emotions

The dreamer experiences three-quarters of the emotions

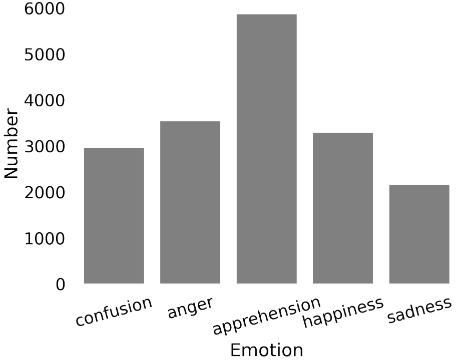
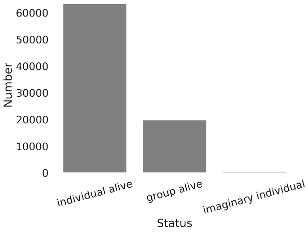
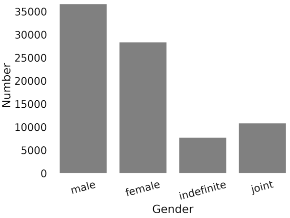


Figure: Distribution of emotional states.

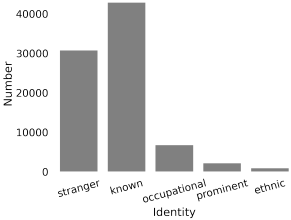
Appendix: distribution of character classes



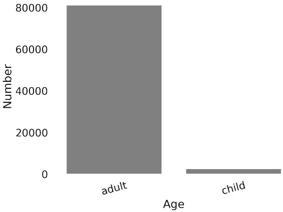
(a) Status



(b) Gender



(c) Identity



(d) Age

Appendix: prompt for StableBeluga

We perform in-context learning with StableBeluga, a 7B decoder-only model, using the following prompt:

System: You are StableBeluga, an AI that follows instructions extremely well. Help as much as you can. You know the Hall and Van de Castle annotation scheme.

User: Classify CHARACTERS (status, gender, identity, and age) in a DREAM REPORT.

Given a DREAM REPORT, you must follow the format: CHARACTERS: [CHARACTER]status is <status>, gender is <gender>, identity is <identity>, age is <age>

Where:

<status> must be in {"1":"individual alive", "2":"group alive", "3":"dead individual", "4": "dead group", "5":"imaginary individual", "6":"imaginary group", "7": "original form", "8":"changed form"}
<gender> must be in {"M":"male", "F":"female", "J": "joint", "I":"indefinite"}
<age> must be in {"A":"adult", "C":"child"}
<identity> must be in {"K":"known", "P":"prominent", "O":"occupational", "E":"ethnic", "S": "stranger"}

Use [CHARACTERS] to separate multiple characters. Do not classify the dreamer.

Assistant:

Bibliography I

[1] Rosalind Cartwright. "Chapter 45 - Dreaming as a Mood Regulation System". In: *Principles and Practice of Sleep Medicine (Fourth Edition)*. Ed. by Meir H. Kryger, Thomas Roth, and William C. Dement. Fourth Edition. Philadelphia: W.B. Saunders, 2005, pp. 565–572. ISBN: 978-0-7216-0797-9. DOI: https://doi.org/10.1016/B0-72-160797-7/50052-5. URL: https://www.sciencedirect.com/science/article/pii/B0721607977500525.

[2] Francis Crick and Graeme Mitchison. "REM sleep and neural nets". In: *Behavioural Brain Research* 69.1-2 (July 1995), pp. 147–155. DOI: 10.1016/0166-4328(95)00006-f. URL: https://doi.org/10.1016/0166-4328(95)00006-f.

[3] Francis Crick and Graeme Mitchison. "The function of dream sleep". In: *Nature* 304.5922 (July 1983), pp. 111–114. DOI: 10.1038/304111a0. URL: https://doi.org/10.1038/304111a0.

Bibliography II

[4] G. William Domhoff. *The scientific study of dreams: Neural networks, cognitive development, and content analysis.* American Psychological Association, 2003. DOI: 10.1037/10463-000. URL: https://doi.org/10.1037/10463-000.

[5] G. William Domhoff and Adam Schneider. "Studying dream content using the archive and search engine on DreamBank.net". In: *Consciousness and Cognition* 17.4 (Dec. 2008), pp. 1238–1247. DOI: 10.1016/j.concog.2008.06.010. URL: https://doi.org/10.1016/j.concog.2008.06.010.

[6] Sigmund Freud. "The interpretation of dreams". In: *Literature and Psychoanalysis*. Columbia University Press, 1983, pp. 29–33.

[7] Michael Schredl and Friedrich Hofmann. "Continuity between waking activities and dream activities". In: *Consciousness and Cognition* 12.2 (June 2003), pp. 298–308. DOI: 10.1016/s1053-8100(02)00072-7. URL: https://doi.org/10.1016/s1053-8100(02)00072-7.

Bibliography III

[8] Serge Thill and Henrik Svensson. "The inception of simulation: a hypothesis for the role of dreams in young children". In: *Proceedings of the Annual Meeting of the Cognitive Science Society*. Vol. 33. 2011.

[9] Matthew P. Walker and Els van der Helm. "Overnight therapy? The role of sleep in emotional brain processing.". In: *Psychological Bulletin* 135.5 (2009), pp. 731–748. DOI: 10.1037/a0016570. URL: https://doi.org/10.1037/a0016570.

[10] Carolyn Winget and Milton Kramer. *Dimensions of dreams*. Gainesville: University of Florida Press, 1979.

[11] Michel Zlotowicz. "Sur l'analyse du cauchemar enfantin". In: *Bulletin de psychologie* 26.305 (1973), pp. 615–621. URL: https://www.persee.fr/doc/bupsy_0007-4403_1973_num_26_305_10419.