SOCIAL SCIENCE REVIEW ARCHIVES
ISSN Online: 3006-4708
ISSN Print: 3006-4694
https://policyjournalofms.com

Beyond the Lexicon: Semantic Approaches to Coherence and Meaning-Making in AI-Generated Texts

Muhammad Zeeshan[1], Anzala Arif[2], Eman Imran[3]

[1]Head of English, PGC. Sialkot
[2]Government College University, Faisalabad
[3]Government College University, Faisalabad

DOI: https://doi.org/10.70670/sra.v4i1.1737

Abstract
This research analyzes the significance of coherent and meaningful AI-generated content in semiotics. Although transformer LLMs provide extremely eloquent results, or even they often times unable to maintain the coherency of discussion, elude false or misleading information, or conserve assertion. The approach introduces a multi-methodological analysis and produce AI-created content by illustrating on analytical approaches, operational models of local coherence(entity-grid) ,contemporary semantic scorers (BERT Score, NLI detectors),and (Rhetorical Structure Theory; attention/intentional accounts).It blueprints a realistic decorum that connects the various checks like entity, persistency, hallucination detectors, human connotation ,discourse relation analysis, and entailment checks to identify where semiotic failures happen and design model advancements. This research asserts that to get AI-generated text and powerful, we need to assimilate semiotic limitations to how we train, decode, and evaluate models–surpassing the vocabulary to knowledge about discourse, contextually rooted in AI writing systems.

Introduction
Large neural language models (LLMs) have drastically enhanced how the nature of the fluency of machine transformed the text: the results are more diverse, realistic and contingently likely than former systems. The literacy is primarily the outcomes of the neural network framework and the ordered series of current training, that commonly allow models to pick up complicated statistical findings in communication. (Aswanet al.; Brown et al.). However, a persistent criticism remains, fluency does not equivalent to meaning. AI generated text fail to maintain strong logical consistent connected meaning throughout a document or article but can be seen as consistent at the sentence. Usually reported issues include logical gaps or non-sense beyond sentences, repetition and degeneration breakdowns in interpretation methods and affirmations that are assured but groundless usually named as “hallucinations” (i.e., fluent but false or unsupported claims) (Holtzman et al.; Maynez et al.; Ji et al.). For a specific semantic coherency and meaning-making, this paper suggests us to enhancing AI-generated content through moving above the language and besides focusing on n-gram fluency and word possibilities. Through linguistics and computational models, I combine discourse and coherence theory for consistency, condense observational data of recent text generation systems and introduce a practical way to evaluate and improve systems that focuses on discourse patterns, semiotic relations and realistic evaluation. I suggest relevant metrics and methods like (entity-grid models, semantic similarity measures, NLI based faithfulness checks, neural semantic Scorers) that can be merged into a semantic processing pipeline for analysis and development.

Theoretical framework: semantics, coherence, and discourse

A lexical method to coherent and meaning-making pulls from multiple identified areas in linguistic and statistical analysis.

Discourse structure, coherence relations, and intentional structure

Classical concepts of the text structure underscores that consistency is higher than proximity. Instead, coherence is contingent on based functional and intentional relationship which connects segments and clauses. Rhetorical Structure Theory (RST) views layouts as a system of hierarchical relations structured by rhetorical function (Man and Thompson). Grosz and Snider's theory underscores intentions and intentional focus as pivotal to discourse organization (Grosz and Sidner). These above-mentioned systems elucidate why superficial structure of sentences can still fail to add up as a meaningful, integrated communication act.

Coherence as entity continuity and semantic relations

The *entity-grid* approach is an important method of local coherency in text. It studies at how these remarkable entities (noun phrase) show up and change through sentences. Coherent text manifest specific patterns of entity distribution (Barzilay and Lapata). Collaborative work highlights rhetorical relations (e.g., causal, contrastive, elaboration) and how their appearance or absence forms analytical connections which reader develop (Hobbs; Mann and Thompson). This understanding tells us to check if AI generated text adheres the same hypothesis as human generated text do far audiences, we must examine how entities continue to evaluate through the text and cohesive devices.

Semantic similarity, entailment, and meaning-preservation

Creating-meaning just not involves in interconnectedness but also maintaining the intentional meaning of the propositions. Modern neural Scorers (e.g., BERT Score), semantic similarity models (contextual embedding's); Natural Language Inference (NLI) models evaluate semantic similarity and inference. These frameworks enable to poses either a generated text keeps the key concepts from the original text instead of simply reiterating surface-level patterns. (Zhang et al.; Maynez et al.).

AI-generated texts and semantic challenges

This segment studies how modern LLMs both enhanced or exposed new semantic issues in AI-generated text.

Why current LLMs produce fluent but sometimes semantically weak text

LLMs use transformer architecture and large-scale pertaining to generate models that make consistencies patterns over a vast amount of text, facilitating situational fulfillments and stylistics imitations (Vaswani et al.; Brown et al.). Yet these frameworks primarily focusing on Next-Token Prediction (NTP), rather than creating unclear communication blueprints, achieving logical goals over an expanded work, or defending the validity of a proposition. Consequently, the findings can be enhanced the local modeling inference although fail to achieve broader semantic restrictions. The gap between the discourse level goals and the research objective enables to analysis several flaws identified in method.

Decoding pathologies: degeneration and surface repetition

This study has verified interpreting issues that lower the quality of content generation maximum likelihood estimation (beam search, greedy) incline to generate monotonous or repetition content; although simple selection may shift into inconsistency. Holtzman et al. indicates that deciphering approach significantly influences the textual richness or cohesion, prompting the strategies like token

Hallucinations, factuality, and faithfulness

Hallucination is the main issue of contextual meaning: paradigms create proposition with confidence unwarranted input data or reality. Hallucination weakens trust and usefulness (Maynez et al.; Ji et al.), for example in study of abstract and free-form text generation. This study shows hallucination is prevailing through are different created tasks and is a major area of study. To address this, researchers generally entail data correction, combining modeling and evaluation models which is directly focus on faithfulness.

Semantic drift and pragmatic incoherence

Outputs can show semantic drift yet they are not actually incorrect. The text gradually takes away from the primary subject or fails to preserve intended assumptions and implications. In long-form text the outcomes in paragraph which is cease to fulfill reader assumptions identified by earlier sentences. It is more about failing to create meaning instead of grammatical problems. Factual accuracy metrics which is disregard semantic relationships away from the context (e.g., surface n-gram overlap), usually overlook this drift. They generate deceptively high scores for apparent fluency but produce, contextually incoherent narratives.

Methodology: a semantic analysis and evaluation protocol

I propose a protocol that mixed quantitative-qualitative methods to study coherence and meaningful AI outcomes which can be applied to candidate corpora (human texts, outputs from different decoders/models).

Data and experimental setup

1Corpora:

Gather similar sets of texts including (a) outputs from multiple LLMs (e.g., GPT-family models, smaller transformer baselines) (b) human-written documents using various (diverse decoding methodologies (greedy, beam, top-k, nucleus). Combine tasks conditional generation (summarization, long-form explanation) and prompt-based to ascertain varied discourse requirements. When portraying on LLM options, cite references for blueprints are essential.

Annotation:

Enroll annotators to evaluate (i) global topic adherence; (ii) factual faithfulness (where relevant); (iii) local sentence-to-sentence coherence and (iv) subjective relevance (does the text convey a clear, interpretable message?). Annotations protocols specify each variable and apply multiple rating through inter scorer consistency reviews.

Automatic metrics (semantic-centered)

Connect multiple automatic metrics instead of depending on just one measure:

- **Entity-grid / local coherence models:**

Determine neural coherence scores (Barzilay and Lapata; neural extensions) and entity transition features. These compute either important entity is preserved and referred to in appropriate models that suit the discourse.

- **Semantic similarity & entailment:**

Apply NLI models and textual-based metrics (BERT score) to evaluate either a created text maintains or changes propositions from the prompt (Zhang et al.; Maynez et al.). Natural Language Inference can spot dichotomies or contradiction that show semantic shift or illusion.

- **Factuality / hallucination detectors:**

Use factual consistency checks (QA-based, IE-based, or LM-based detectors) as examined in hallucinations surveys (Ji et al.). The certain methods find unsupported evidence and quantify dominance of hallucinated data.

- **Discourse relation parsers:**

If available, use RST or discourse relation parsers to identify unknown and ill-logical rhetorical relations (Mann and If Thompson). Relating rhetorical relation studies with entity continuity suggests a comprehensive analysis on where a text fails to create sense.

Qualitative analysis

For a close reading, choose a classified portion of AI-generating texts, (Evaluate how meaning expands how theories are presented, verified, and concluded also detect basic error types (In discourse theory (anchor the qualitative classification for rigorous interpretation.

Evaluation protocol and fusion scoring

Address the relationship between human judgments and automatic metrics to estimate standard credibility. Apply a combined lexical coherence created through a score fusion of hallucination detectors, BERT Score/entailment signals, and entity-grid features; and adjust the weights through the calibration on human judgment scores. These combined score offers practical, linguistically based unified measure for evaluating models.

Analysis and (interpreted) findings

Subsequently, I propose a narrative analysis dependent on the framework; above-mentioned; as the paper, fails to disclose unknown corpus-based empirical research, the outcomes can be structured like predicted patterns or descriptive analytics based on previous studies.

Expected correlations: fluency vs. semantic coherence

Earlier studies purpose that the apparent proficiency broadly inter-relates, suggesting a semantic measures. The extreme intricacy and significant co- occurrence metrics fail to ensure inference and discourse coherence (Holtzman et al.; Maynez et al.). Therefore, the researchers predict automated fluency measures (LM likelihood, per laxity) to elucidate merely a part of subjective relevance. Although, semantics similarity (entity continuity+ entailment) inter-relates deeply with assessments relating to either an entire text "makes sense" overall.

Typical error typology (illustrative)

AI-generate text often reoccurs the following error types based on surveys and analytical studies:

- **Local referential breakdowns:** specific terms or pronouns with ambiguous references, creating transient confusion identified through referent resolution errors (Barzilay and Lapata) entity-grid shifts.
- **Topic drift:** incremental semantic shift in original text; auto-detection occurs due to declining topical resemblance and abandonment of inference scorers throughout the paper.
- **Propositions Hallucinated:** statements unsubstantiated by common knowledge or by the disclosed via QA-based checks and hallucination detectors (Ji et al.; Maynez et al.).

- **Contradictions and circularity:** Natural Language Inference can identify direct inconsistencies. Models generate reciprocally incompatible claims in the same context.

Case illustration (hypothetical)

Examine a stimulus seeking an abstract of research papers regarding experimental study. A text generation model might generation a section which is coherent yet introduce unsupported effectiveness assertion (e.g., "the drug cured 80% of patients") doesn't report in data. The inference validation illustrates limited connection with data reports. The Question-Answer Extractors lacks to identify substantiating statement in the study; coders highlight the assertion like fabricated. Coherence grid variables can continue seemingly reasonable (same entity names reused), yet logical inconsistency is meaninglessness which is entirely superficial methods (Maynez et al.; Ji et al.).

Metric performance expectations

Research-based texts indicate which integrating-grounded measures for example BERT score connect more effective with individual evaluation compared to clear n-gram frequency, Although QA checks or task-specific NLI offer precise identification of illusion hallucinations (Zhang et al.; Maynez et al.). Integrating such metrics through consistency framework generates better co-ordination using human evaluation of relevance or consistency in the research examined.

Discussion

Why semantics must be central to generation

The core concept is systematic: whether the purpose of creation is to generate significant content which helps subjective understanding, assessment, learning goals must needs to execute semiotic restrictions more than just a symbolic indication. Research framework or analysts enhance proficiency, yet they lack to assure discourse purpose, practical relevance and statement maintenance. Discourse models (RST, Grosz, and Sidner) substitute an ideal objective which is similar to consistent interpretation, although object or implication metrics offer practical indicators for improvement or assessment.

Practical model interventions

Various general guidelines pursue organically:

- **Semantically informed training objectives:** integrate supportive casualties which indicate an incentive to original material, entity continuity, and estimating connection throughout optimization. Preliminary studies on restricted aims or multi-objective optimization proposes practicality.
- **Hybrid symbolic-neural pipelines:** integrate neural production with formal verifications (fact-checkers, knowledge graph inference, NLI filters) implemented in interpretation and output processing to identify or accurate the semiotic mistakes. Questionnaire on false perception reduction emphasize the guarantee of integrated systems.

Evaluation and research practice

Scholars or experts must account not only proficiency yet hybrid human evaluations or meaningful consistency ratings centered on interpretations. Evaluation packages must incorporate activities which is directly assess communicative purpose (e.g., argument continuation, explanation coherence) and discourage unsubstantiated statement. It correlates paradigm evaluation to actual research objectives of content creation.

Implications and future directions

For NLP research:

- Create metrics or collaborate assignments which classify semantic preservation (e.g., datasets with entailment-annotated summaries, long-range coherence texts) and textual cohesion.
- Enhance coherent neural framework which combine object, consistency rhetorical links or syntactic logical form united result.

For industry and deployment

Combine meaningful assessments for the creation of workflows in a core aspect, (health, laws, finance), at which inconsistency and hallucinations possess actual outlays. Apply NLI classifiers, structured knowledge validation like minimum security measures (Ji et al.; Maynez et al.) QA-based checks.

For ethics and user trust

- Openness relating to semiotics restrictions or vigorous UI that indicate ambiguous or unsupported ascertains which is vital to participant trustworthiness. Frameworks have to manifest trust and origins, more than refined writing.

Conclusion

Enhancing auto-created content to surpassing verbal proficiency toward a more affluent, contextual interpretation. Research instrument via Natural Language Processing (NLP) (entity grid, discourse parsers), models via text analysis (RST, intentional structure) or modern meaning analysis (BERT score, NLI detectors) collectively from a applicable framework for the purpose of detecting or optimizing meaningful consistency within created content. Main objectives of the area of study encompasses 1) Educational targets and interpreters which integrate meaning restrictions, 2) Implementation process which protect hallucination by means of validation, and 3) Developing assessment measures that represent the creation of meaning. Whether the purpose of AI content generator is genuinely to convey thoughts on which users can trust and take action. Therefore contextual-beyond the terminology should be focused on methodology or assessment.

Reference

Barzilay, Regina, and Mirella Lapata. "Modeling Local Coherence: An Entity-Based Approach." *Computational Linguistics*, vol. 34, no. 1, 2008, pp. 1–34. ACL Anthology.

Brown, Tom B., et al. "Language Models Are Few-Shot Learners." *Advances in Neural Information Processing Systems*, 2020. NeurIPS Proceedings, https://arxiv.org/abs/2005.14165.

Grosz, Barbara J., and Candace L. Sidner. "Attention, Intentions, and the Structure of Discourse." *Computational Linguistics*, vol. 12, no. 3, 1986, pp. 175–204. ACL Anthology.

Hobbs, Jerry R. "Coherence and Coreference." *Cognitive Science*, vol. 3, no. 1, 1979, pp. 67–90.

Holtzman, Ari, et al. "The Curious Case of Neural Text Degeneration." arXiv, 2019.

Ji, Ziwei, et al. "Survey of Hallucination in Natural Language Generation." *ACM Computing Surveys / arXiv*, 2022 (rev. 2024), arXiv:2202.

Mann, William C., and Sandra A. Thompson. "Rhetorical Structure Theory: Toward a Functional Theory of Text Organization." *Text — Interdisciplinary Journal for the Study of Discourse*, vol. 8, no. 3, 1988, pp. 243–281.

Maynez, Joshua, Shashi Narayan, Bernd Bohnet, and Ryan McDonald. "On Faithfulness and Factuality in Abstractive Summarization." *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL)*, 2020, pp. 1906–1919.

Vaswani, Ashish, et al. "Attention Is All You Need." *Advances in Neural Information Processing Systems (NeurIPS)*, 2017, https://arxiv.org/abs/1706.03762.

Zhang, Tianyi, Varsha Kishore, Felix Wu, Kilian Q. Weinberger, and YoavArtzi. "BERTScore: Evaluating Text Generation with BERT." arXiv, 2019 (ICLR 2020).