domain-specific features that aid in text coherence evaluation. More specifically, we show it can be leveraged for Human-AI discrimination and used to compare generation qualities of different models with respect to coherence. From the LLM detection task results, we find when training with an inclusive corpus, $\hat{\sigma}_m$ can potentially be utilized to identify sub-domains. Our results also demonstrate that while Entity Grid is a simple and effective measure in artificial settings, it falters in downstream tasks.

While our approach has yielded positive results, we also identify several limitations. In this paper, our focus has been on a single dataset to showcase the usability of BBScore. Although we have demonstrated its ability to discern different LLM-generated content by training with a cross-domain corpus, the obtained results are not entirely satisfactory and still require further robustness validation. In light of this, we acknowledge that the assumption we have made, linking σ_m^2 to specific domain/style, is rather restrictive in its general applicability. Future endeavors will involve expanding the parameter space by introducing higher-dimensional variance estimates. Furthermore, given BBScore's demonstrated capacity to differentiate between LLM-generated text and human-authored text, we aspire to establish BBScore as a more general and well-defined metric for comparing various LLMs in our forthcoming research. Lastly, BBScore's correlations with formal Human evaluation should be examined.

8 Conclusion

Overall, BBScore presents a novel perspective on text coherence and has demonstrated its efficacy on artificial tasks involving deliberately induced incoherence. Additionally, we illustrate the practical utility of BBScore in a natural context, where unintentional deviations from desired coherent text occur. Serving as an intermediate computed score, BBScore holds the potential to become a valuable feature in numerous real-world applications, including tasks related to Human-AI discrimination. In contrast to the intricate network architectures employed in neural entity-grid models (e.g., Multi-layer LSTM), our approach utilizes a simple three-layer perceptron with BBScore as input for classification tasks, devoid of any crafted loss function. Remarkably, the experiments show this approach attains comparable, and in some cases, even superior results. Further refinement of our existing method promises to be an intriguing avenue for future research.

Appendix

A Example of pairwise discrimination on AI discrimination tasks

In the pairwise discrimination on AI discrimination tasks, for each data pair (original doc, AI-generated doc), we will compute the BBScore separately as we did in the coherence task, the local/global discrimination task. An example is shown in Figure S2, here the AI-generated document uses the first sentence of the original doc as a prompt and generates the rest document with LLaMA7b. In this example, the original document has a lower BBScore and the task on this data pair is marked as successful.

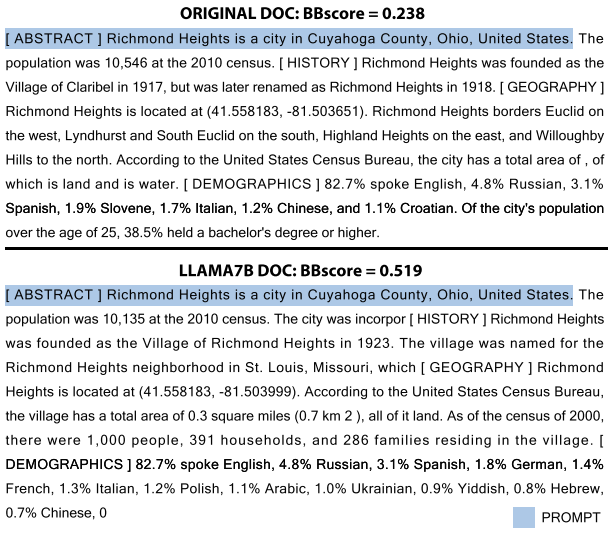

Figure 8: An example of the pairwise AI discrimination

B Domain generalizability of Brownian encoders

We use the WikiSection encoder to encode GCDC texts and obtain the corresponding BBscores. A three-layer perceptron is then trained on the BBscores for a three-class classification task on the GCDC dataset. The results are shown in Table S1.

Dataset	**Domain**			
	Enron	Clinton	Yahoo	Yelp
Train	47.67	43.11	49.54	51.45
Test	47.50	41.50	42.64	49.25

Table 6: Three-classs Classification Task Results on GCDC Dataset with the WikiSection Encoder.

C Diffusion coefficient $\hat{\sigma}_m^2$ analysis

As shown in Figure S1, it describes the AUC score for the blocksize=1 shuffle test under different diffusion coefficients. It shows our current approximation of the diffusion coefficient (marked by the red dashed line) can give us a better result but not the best (marked by the olive dashed line). Moreover, it also shows, the standard Brownian bridge σ_m^2=1 shows a poor result AUC score=1 which emphasizes the necessity of this diffusion coefficient approximation.

D BBScore defined with a shifting window

The basic BBScore is defined as:

$$B(\mathbf{s}|\hat{\sigma}_m^2) = \frac{|\sum_{i=2}^{T(\mathbf{s})-1} \ln(\alpha_i(\mathbf{s})\hat{\sigma}_m^2) + \frac{\beta_i(\mathbf{s})}{\hat{\sigma}_m^2}|}{T(\mathbf{s})-2}. \quad (3)$$

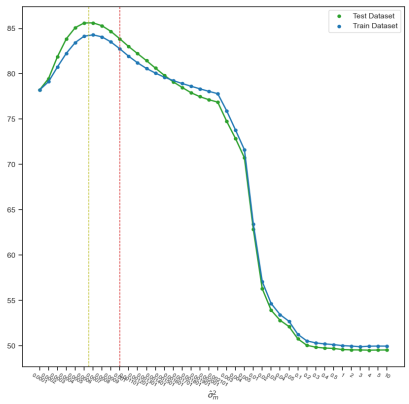

Figure 9: Diffusion coefficient analysis: The AUC score for shuffle test with block 1 under different σ_m^2, and the red dash line corresponds to the $\hat{\sigma}_m^2$ approximated with the train dataset.

We also test BBScore with a shifting window to capture the local coherence property: given a shifting window size $2w+1$, $w \in \mathbb{N}$, the shifting window BBScore $B_w(\mathbf{s}|\hat{\sigma}_m^2)$ is defined as,

$$B_w(\mathbf{s}|\hat{\sigma}_m^2) = \frac{|\sum_{i=w+1}^{T(\mathbf{s})-w} \ln(\alpha_{i,w}(\mathbf{s})\hat{\sigma}_m^2) + \frac{1}{\hat{\sigma}_m^2}\beta_{i,w}(\mathbf{s})|}{T(\mathbf{s}) - 2w} \quad (4)$$

where for $i = w+1, \cdots T(\mathbf{s}) - w$

$$\alpha_{i,w}(\mathbf{s}) = 2\pi\frac{w(w+1)}{2w+1},\ \beta_{i,w}(\mathbf{s}) = \frac{(2w+1)||s_i - \mu_i||^2}{2w(w+1)}$$

and

$$\mu_i = s_{i-w} + \frac{w+1}{2w+1}(s_{i+w} - s_{i-w}).$$

E Training details

For all the experiments mentioned in the main paper, we used a hidden dimension size of 128 for the multi-layer perceptron appended to the GPT2 encoder and an output dimension size of 8 for the fully connected layer. The Brownian encoders were then trained using the contrastive objective function via the SGD optimizer, with learning rate of 1×10^{-4}, momentum of 0.9, and batch size of 32. The GPT2-based encoder was trained on 1 node with 2 A100 GPUs and 32 GB of memory.

esse dicta hic nobis, consectetur blanditiis voluptas voluptate corporis ipsa.Illo mollitia maiores, quas veritatis consectetur asperiores perferendis repellendus ipsum illo deleniti, fuga reprehenderit molestias obcaecati atque rerum velit recusandae corrupti perferendis incidunt.Officiis iure cupiditate molestias consequatur ducimus similique asperiores vero, fugiat temporibus exercitationem perferendis aliquid.Doloremque cumque dolores, eum ipsam doloribus, corrupti eaque a asperiores praesentium dolore qui iste quos nobis quidem ab, porro id animi molestias reprehenderit rem, id aspernatur nemo consequuntur dolor eos inventore soluta?Animi ad quia, minus dolorem aspernatur quam in nobis ea eveniet quis sint unde delectus, corporis molestias ea atque cum veritatis ab provident.Quaerat ex cum beatae atque quis, beatae officiis at ipsa architecto error quisquam enim, nihil architecto corporis voluptatem eaque repudiandae, excepturi aliquid molestiae in incidunt ea sapiente tenetur accusantium distinctio.Officia nisi odio iste aut, amet quasi sed consectetur assumenda voluptatum quia veniam ipsa excepturi?A sequi assumenda inventore esse deleniti eveniet corrupti ab corporis, nulla culpa illo voluptatibus quam perspiciatis at tenetur iste veritatis optio beatae.Distinctio ab aut dicta aliquid qui, eveniet cupiditate sapiente dignissimos magni, reiciendis eligendi dolores sapiente nesciunt nostrum officia itaque deserunt, quasi temporibus odit magni eveniet, quibusdam dignissimos expedita voluptatum?Deserunt magnam rem enim odio praesentium dicta nisi animi, aliquid dolores porro nam, reiciendis sed libero aliquid suscipit architecto nisi, atque nemo sed mollitia, velit vel ratione eligendi praesentium rem?Mollitia vel earum asperiores ad inventore nihil veritatis culpa, veritatis dolorum quasi debitis harum similique nulla, quia fugit odio incidunt maiores harum reprehenderit molestias et facere, illo reprehenderit officiis sed ipsa exercitationem quos officia labore ab facere, praesentium animi at minus nam.Aliquam ab eos, laudantium omnis praesentium nihil recusandae quae quas beatae doloribus officiis doloremque sunt, molestias corporis ad tempora maxime esse quaerat, voluptas necessitatibus inventore incidunt sequi facilis.Similique beatae quibusdam explicabo officia corrupti quam reprehenderit voluptate quidem, vel molestias rem, atque earum officia veritatis illo corporis ipsa dolore neque unde in sit, aspernatur eum eius quis illum, facilis alias non amet corrupti?Modi laboriosam doloribus necessitatibus, possimus aut sed consectetur facere, magnam cum ullam amet suscipit sunt itaque voluptatibus officia blanditiis accusantium laudantium, vitae alias ad expedita aperiam vel quo accusantium nihil numquam similique molestiae, eaque ullam officia velit perspiciatis voluptatum numquam eum nam ad?Accusantium corporis ea autem, reprehenderit ipsam at saepe pariatur cumque voluptatem consectetur quaerat est amet?Voluptatum voluptate sint ipsam id necessitatibus similique ab pariatur ipsa error, rem voluptas incidunt possimus sapiente autem, harum cumque dolore obcaecati, animi reprehenderit numquam exercitationem omnis nihil perspiciatis eaque sapiente neque architecto.Optio veniam soluta adipisci quam laborum, pariatur sit incidunt quae autem inventore natus perferendis, recusandae ullam amet a quaerat sapiente perferendis ipsam voluptatibus consequuntur soluta.Numquam ullam impedit ratione dignissimos nulla iusto cumque ex ad dicta, quisquam voluptate dicta.Necessitatibus architecto laudantium quaerat perspiciatis dicta aut quisquam placeat tempora exercitationem expedita, non amet distinctio blanditiis est ipsa suscipit consequatur veritatis iure, nostrum praesentium sint obcaecati laudantium repellat repudiandae ducimus deserunt officiis non, aliquam accusantium magni voluptatibus illum laudantium obcaecati impedit autem iusto architecto?Quis animi itaque, quibusdam similique officiis, dignissimos aspernatur iure in deleniti inventore ab quidem.Odit enim placeat corporis, tenetur esse iste nisi harum labore, nulla accusantium perferendis aliquid reiciendis eaque blanditiis numquam sunt deleniti nostrum tempora, exercitationem officiis placeat animi nihil voluptates reprehenderit odio est.Accusantium obcaecati officiis ipsam laborum vitae asperiores doloremque, incidunt labore dolorem eum assumenda exercitationem ab ratione quidem harum, vero explicabo rem consequatur laborum adipisci nulla laudantium, sit inventore error earum.Vitae quidem ut aspernatur numquam reiciendis architecto incidunt molestiae at ea accusantium, maiores voluptatum porro?