Blabrecs: An AI-Based Game of Nonsense Word Creation

Max Kreminski
Santa Clara University
mkreminski@scu.edu

Isaac Karth
Liminal Experiences, Inc.
isaac@liminal.gg

Abstract

Blabrecs is an AI-based game of nonsense word creation that modifies the popular wordgame *Scrabble*, substituting the English dictionary that's usually employed to adjudicate word validity with an AI-based classifier. This classifier has been trained to accept only letter sequences that sufficiently *resemble* English words; actual English dictionary words are disallowed from play, so in order to achieve victory, players must invent nonsense words that closely resemble English. The result is a play experience that promotes and celebrates a diverse array of linguistic innovations by players, rather than homogenizing wordplay around a standard set of "good *Scrabble* words" as ordinary dictionary-based *Scrabble* gameplay tends to do. This paper briefly describes *Blabrecs* in the context of the NeurIPS 2023 Creative AI exhibition, at which it will be demonstrated.

1 Description

Blabrecs [8] is a hybrid digital/physical board game and a rules modification to the popular wordgame *Scrabble*. In *Blabrecs*, as in *Scrabble*, players take turns drawing letter tiles from a bag and placing these tiles on a grid to form words, which are then scored based on letter frequencies and tile score multipliers to award players with points. Unlike *Scrabble*, however, *Blabrecs* does not use an English dictionary to determine what letter sequences constitute valid words. Instead, it uses a classifier *trained on* the English dictionary to accept or reject letter sequences. Actual dictionary words are disallowed; only nonsense sequences that the classifier misclassifies as words are allowed to be played. Consequently, *Blabrecs* players are forced by the game's constraints to innovate new words that closely resemble English (enough to fool the machine) but that are not already contained in the "standard", dictionary-defined vocabulary of the English language.

In our exhibition of *Blabrecs* at NeurIPS 2023, we will set up one or more tables with *Scrabble* sets for conference attendees to use in playing *Blabrecs* together. Attendees will be able to walk to the *Blabrecs* exhibit space, sit together around a table, and play *Scrabble* while using the *Blabrecs* AI classifier to determine which nonsense words are allowable for play. The *Blabrecs* web interface will be accessed directly by attendees on their own devices, for instance their mobile phones. This interface will automatically accumulate a list of all nonsense words that actually get played; this list includes a text box next to each word that the players can freely edit to give each word a custom definition. At the end of a play session, players can take home a copy of this list of words and definitions, as well as photographs of the board state, as a record of the words they invented together.

Though *Blabrecs* does not ordinarily record player-submitted words and definitions past the end of a single play session, we plan to temporarily enable opt-out logging of *Blabrecs* words during the NeurIPS 2023 Creative AI exhibition period so that we can curate and highlight words that are invented by conference attendees. This will enable us to showcase the diversity of linguistic innovations produced by attendees. Curated words will be made available on the *Blabrecs* website.

37th Conference on Neural Information Processing Systems (NeurIPS 2023).

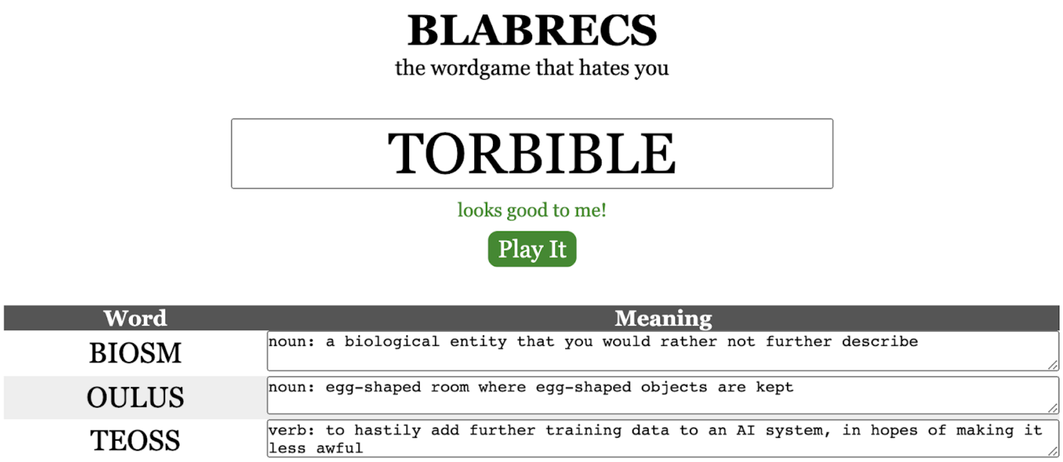

Figure 1: A screenshot of the *Blabrecs* web interface. Above: a text box into which the player types a nonsense word, then receives a judgment from the AI classifier as to whether this word is sufficiently similar to English to be allowed. Below: a table of previously accepted words, along with a player-editable definition for each word.

Both attendees and non-attendees of the NeurIPS conference can access the *Blabrecs* web interface (Fig. 1) online at `https://mkremins.github.io/blabrecs`.

2 Implementation

Physically, *Blabrecs* consists of a standard *Scrabble* set plus a digital device running the *Blabrecs* web interface, which is a clientside-only web app written in HTML, CSS, and ClojureScript. The AI component of *Blabrecs* consists of two classifiers: a Markov chain-based classifier and a more sophisticated classifier based on a separable convolutional neural network, both of which run directly in the web browser. Players can freely switch between these two classifiers as they play.

2.1 Markov Chain Classifier

The initial implementation of the *Blabrecs* classifier is based on a Markov chain trained on the ENABLE word list[1], which is often used as a baseline English dictionary for word games. To train the model, we first turn each word in the word list into a sequence of character trigrams; for instance, the word "apple" is turned into the sequence `["^ap", "app", "ppl", "ple", "le$"]` (where the ^ and $ characters represent the start and end of a word respectively). Then we calculate and store the frequency of each trigram relative to other trigrams that begin with the same two-character prefix.

To evaluate the plausibility of a letter sequence using this model, we divide it into a sequence of trigrams as before and look up the frequency of each trigram in the Markov chain. The per-trigram frequencies are first multiplied together to determine an overall likelihood score for the input letter sequence; this score will always be 0 if the sequence contains any trigrams that were not present within the ENABLE word list, and longer sequences will generally produce lower scores. Then we check whether this score is above or below the *average* likelihood score for real dictionary words of the same length. If the letter sequence is both more likely than the average real word of this length and does not appear in the dictionary, we allow it to be played (Fig. 2).

This classifier is quirky. In particular, it can often be convinced to accept words that contain some highly implausible trigrams if several highly plausible trigrams are also present. Additionally, with the exception of the first and last trigram in each word, it pays no attention to where in the word a trigram occurs. Nonetheless, this classifier was the only one present when the game was first launched, and it seems to mirror the typical player's intuitive sense of plausibility well enough to make for interesting play.

[1] `https://www.wordgamedictionary.com/enable`

Figure 2: A visualization of how the Markov chain classifier works: by splitting the player's submitted word into character trigrams, evaluating the likelihood of each trigram, multiplying together these likelihoods, and comparing to a probability threshold derived from actual English dictionary words of the same length.

2.2 Neural Classifier

An alternative implementation of the *Blabrecs* classifier makes use of a separable convolutional neural network (CNN). This classifier is modeled loosely on the CNN-based text classifier presented in Step 4 of the Google Developers text classification guide [5], but modified to work in TensorFlow.js (so that it can be used in a web browser) and to treat characters as tokens instead of words (since our goal, unusually for text classification, is to classify sequences of up to 16 letters, rather than longer passages of text).

Because ENABLE alone proved to contain too little data to train a good CNN, this classifier was instead trained on three word lists: the YAWL[2] (a strict superset of ENABLE), Letterpress[3], and Moby[4] word lists. These word lists were concatenated together, and duplicate words were removed. Additionally, we generated 2,016,000 unique non-word sequences of random letters between 3 and 24 letters in length to use as negative examples; this is approximately six times as many negative examples as there are positive examples in the combined word list.

For the negative example generation task, we used a weighted random process to select letters at the same rate that they appear in known English words (Fig. 3). An earlier attempt at training a classifier with unweighted random letter sequences as negative examples resulted in the acceptance of far too many subjectively implausible words. This may be because unweighted random letter sequences are too easy to differentiate from English words based on letter frequencies alone, resulting in poor performance when the resulting classifier is confronted with the types of approximately English-letter-frequency nonsense words that players tend to submit in practice.

To evaluate a player-submitted letter sequence, we use the classifier to predict its likelihood of being a valid English word and check whether the predicted likelihood is greater than 0.82. This threshold was determined by manually testing a large number of words and picking a cutoff that seemed to match our intuitive notion of word plausibility.

Evaluating the quality of a nonsense word gatekeeper is difficult and largely intuition-driven. However, the neural classifier seems to match the authors' intuition for nonsense word plausibility more reliably than the Markov chain classifier; in particular, it seems less prone to "false negatives", or judging nonsense words as implausible that the authors consider plausible. Additionally, the neural classifier's quirks are less obvious and easy to learn than those of the Markov chain classifier: it is more difficult

[2] `https://github.com/elasticdog/yawl`
[3] `https://github.com/lorenbrichter/Words`
[4] `https://www.gutenberg.org/files/3201/files/SINGLE.TXT`

Figure 3: Example training data for the CNN classifier. Positive examples were actual dictionary words; negative examples were random letter sequences with the same letter frequencies as actual English text.

to figure out what features the neural classifier weights most strongly in its estimation of nonsense word plausibility.

2.3 Limitations and Future Work

The Markov chain and CNN-based methods of classification were selected at the time of *Blabrecs*'s original development (from late 2020 to mid-2021), and the state of the art in many AI subfields has evolved substantially since that time. As a result, it is unknown how well these particular classification methods perform on the *Blabrecs* classification task compared to others in the rapidly expanding space of possible approaches.

In the future, we may develop and deploy additional classifier options: for instance, a classifier based on a transformer or LSTM architecture. However, we also require our classifiers to function entirely within the user's web browser to ensure that we do not need to maintain a *Blabrecs* server, so our classifiers are limited in the amount of computational power they can demand.

3 Design

3.1 Design Methodology

To validate our high-level gameplay concept, we began the design process for *Blabrecs* by creating a Wizard of Oz prototype [7, 2] in which a human played the role of the AI, judging letter sequences as valid or invalid on the basis of intuition. Several rounds of playtesting revealed that the invention of feasible nonsense words to bypass a gatekeeper agent could produce a compelling play experience, so we went forward with a computational version of the game.

In its use of AI to support a gameplay experience that wouldn't be straightforwardly possible without an AI component, *Blabrecs* represents a clear example of *AI-based game design*: an approach to game design that strives to make effective use of AI's unique material affordances [4, 19]. Specifically, we view *Blabrecs* as an unusual simultaneous example of the *AI as co-creator* and *AI as adversary* design patterns proposed by Treanor et al. [19]: players co-create with an AI adversary that supports their creativity by imposing otherwise difficult-to-enforce creative constraints.

3.2 Design Goal: Promoting and Celebrating Diversity

One of our primary goals in creating *Blabrecs* was to use AI to promote greater diversity in language use. *Scrabble* as a game is notorious for its imposition of an external authority (the dictionary) between players and their own language [3, 20], and in this way it bears some resemblance to AI-based tools that try to standardize the use of language—sometimes via basic normalization of spelling and grammar, and sometimes even by fully rewriting the user's words into a different and more "correct" writing style, as in some recent LLM-based approaches to helping people write. *Blabrecs* was originally conceived in part as an artistic protest against the standardization of language in all of its many forms.

By imposing deliberately absurdist constraints on language usage, *Blabrecs* forces players to invent entirely new words. The web interface also provides space for players to write in their own definitions for these words, giving players even further support for generating a diverse vocabulary of new words. Based on our observations of past playtests of *Blabrecs*, the kinds of words that players invent are often influenced by their own diverse experiences and backgrounds, and no two player groups are likely to end up repeating the same word. This is a far cry from the standardization of language seen in high-level *Scrabble* play, where certain strategically significant words are memorized and employed in almost every game, despite their extreme rarity in ordinary conversation and writing.

More generally, we hope that the way AI is used in *Blabrecs* (to push players *away* from the typical, rather than pushing them towards it) provides a vision for how AI might be used similarly in other creative contexts. In a time of increasing AI-driven standardization of creative form [1, 10, 14, 11], we feel that the use of AI to specifically *promote* (rather than undermine) creative diversity is an important and underinvestigated direction for research in AI-based creativity support tools.

3.3 Other Design Goals

Beyond promoting and celebrating diversity in language use, we had two other significant design goals for *Blabrecs*. First, we wanted to demonstrate how gameplay could be used to help players develop an intuitive feel for how an AI system works. In the course of *Blabrecs* gameplay, players are strongly incentivized to discover and exploit quirks in the AI gatekeeper's evaluation process; additionally, players can compare and contrast how words are evaluated by two different classifiers. As a result, players may come away from *Blabrecs* with a stronger intuitive sense of how their writing might be evaluated by different kinds of AI systems. Our approach here was influenced by Long et al.'s *AI literacy* framework [12] and subsequent incorporation of this framework into playful educational exhibit design [13].

Second, we wanted to create a game in which players build up a private language with one another as they play. In each *Blabrecs* play session, as players play new words, they are added to a table of player-editable definitions, allowing the players to collectively decide on meanings for the words they have invented. Some of these words may live on within the group of players as in-jokes, mirroring the way that a private lexicon is invented between the players in tabletop language creation games like *Dialect* [18].

4 Ethical Considerations

4.1 Copyright and Licenses

Both versions of the *Blabrecs* classifier were trained only on freely available wordlists that are widely used in other wordgames.

To whatever extent possible, we consider players of *Blabrecs* to be the owners of any nonsense words they generate during the course of play, including any definitions for these words that they enter into the web interface. Except when needed to curate and display submitted words during specific exhibition periods (such as the NeurIPS 2023 Creative AI exhibition period), we don't log or store the words that players invent in any form—the AI components of *Blabrecs* run exclusively clientside in the player's browser. During the exhibition, we will prominently advertise the logging of words and their definitions in the *Blabrecs* web interface and allow players to easily opt out.

4.2 Safety and Security

Though there are relatively few potential safety issues associated with *Blabrecs*, we have implemented a denylist of slurs (based on the `dariusk/wordfilter` module [6]) that we disallow as substrings in *Blabrecs*-approved words. This helps to ensure that the AI classifier doesn't end up appearing to approve of prejudice when it judges words as acceptable for play.

Since the *Blabrecs* web interface doesn't collect any form of data (except for submitted words and their definitions needed to curate words for display during specific exhibition periods, such as the NeurIPS 2023 Creative AI exhibition period), and because players are totally anonymous, there are no other significant security or safety risks to mention. Though players could technically submit personally identifiable information in the open-ended text definitions of invented words via the *Blabrecs* web interface, we will be actively curating submitted words prior to their display in order to screen out any such definitions, and submitted words will be stored securely such that only the *Blabrecs* developers have access to the pre-curation log of submitted words. Additionally, players will be made aware of the logging as soon as they access the *Blabrecs* web interface and will be allowed to easily opt out.

5 Related Work

In addition to the aforementioned language creation tabletop game *Dialect*, several other AI-based language games and explorations served as sources of design inspiration for *Blabrecs*. The *Scrabble*-like word construction game *Rewordable* [15] is of particular note for how the designers made use of AI to identify a set of letter sequences that could be used as cards to improve on *Scrabble*'s letter-tile-based gameplay. Unlike in *Blabrecs*, however, the *Rewordable* player does not interact directly with an AI system.

One of the first author's previous AI-based game projects—*Throwing Bottles at God* [9]—represents an earlier attempt to make Markov chains playable. Rather than classifying player-submitted text, *Throwing Bottles* makes use of Markov chains as a predictive text algorithm to help the player write short messages in a particular style. This can be viewed in hindsight as a failed experiment, whereas *Blabrecs* has been much more successful in eliciting the desired player experience.

The recent art project *New Words* [17] uses the CLIP model [16] to invent appropriate-feeling English-like words for otherwise unnamed concepts directly, rather than acting as a discriminator paired with a human generator of candidate words (in the manner of the *Blabrecs* classifier). *New Words* may be a source of design inspiration for future work on *Blabrecs*, for instance as we investigate additional classifier architectures that we could potentially add to the game.

6 Conclusion

We have presented *Blabrecs*, an AI-based wordgame that subverts the game mechanics of *Scrabble* to promote linguistic diversity rather than homogenization. We hope that the presentation of *Blabrecs* at NeurIPS 2023 will help to encourage future work in the use of AI to promote creative diversity. Readers can play *Blabrecs* at `https://mkremins.github.io/blabrecs`.

Author Biographies

Max Kreminski is a researcher in artificial intelligence, human-computer interaction, and creativity; an assistant professor of computer science and engineering at Santa Clara University; and a research scientist at Midjourney, where they direct the storytelling tools research group. They have published over 50 papers in HCI and creative AI venues ranging from the ACM CHI conference to the International Conference on Computational Creativity. Eight of their papers have won or been nominated for Best Paper or similar awards. Max has been an invited speaker at numerous venues, including the industry-leading Game Developers Conference (GDC), and an AI art toolmaker-in-residence at Stochastic Labs; their work and science communication has been covered by outlets including NPR, New Scientist, Bloomberg, The Verge, The Next Web, and Polygon. Max holds a PhD in Computational Media from the University of California, Santa Cruz.

Isaac Karth is a researcher in artificial intelligence and game design, with a particular focus on procedural generation. He currently works at Liminal Experiences, Inc., creating procedural generation systems that enable non-programmers to create compelling experiences. His academic work has been published at venues including FDG, AIIDE, and DiGRA. Isaac holds an MFA in Arts and Technology from the University of Texas at Dallas and a PhD in Computational Media from the University of California, Santa Cruz.

References

[1] Kenneth C Arnold, Krysta Chauncey, and Krzysztof Z Gajos. Predictive text encourages predictable writing. In *Proceedings of the 25th International Conference on Intelligent User Interfaces*, pages 128–138, 2020.

[2] Michel Beaudouin-Lafon and Wendy E Mackay. Prototyping tools and techniques. In *The Human-Computer Interaction Handbook*, pages 1043–1066. CRC Press, 2007.

[3] Charles Bethea. The battle over Scrabble's dictionaries. `https://www.newyorker.com/sports/sporting-scene/the-battle-over-scrabbles-dictionaries`, 2015. Accessed: 2023-10-28.

[4] Mirjam P Eladhari, Anne Sullivan, Gillian Smith, and Josh McCoy. AI-based game design: Enabling new playable experiences. Technical report, UC Santa Cruz Baskin School of Engineering, Santa Cruz, CA, 2011.

[5] Google Developers. Step 4: Build, train, and evaluate your model. `https://developers.google.com/machine-learning/guides/text-classification/step-4`, 2021. Accessed: 2023-10-28.

[6] Darius Kazemi. wordfilter. `https://github.com/dariusk/wordfilter`, 2013. Accessed: 2023-10-28.

[7] John F Kelley. An empirical methodology for writing user-friendly natural language computer applications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 193–196, 1983.

[8] Max Kreminski and Isaac Karth. A demonstration of Blabrecs, an AI-based wordgame. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, volume 17, pages 243–245, 2021.

[9] Max Kreminski and Noah Wardrip-Fruin. Throwing Bottles at God: Predictive text as a game mechanic in an AI-based narrative game. In *International Conference on Interactive Digital Storytelling*, pages 275–279. Springer, 2018.

[10] Max Kreminski, Isaac Karth, Michael Mateas, and Noah Wardrip-Fruin. Evaluating mixed-initiative creative interfaces via expressive range coverage analysis. In *IUI Workshops*, pages 34–45, 2022.

[11] Isabelle Levent and Lila Shroff. The model is the message. In *The Second Workshop on Intelligent and Interactive Writing Assistants*, 2023.

[12] Duri Long and Brian Magerko. What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020.

[13] Duri Long, Takeria Blunt, and Brian Magerko. Co-designing AI literacy exhibits for informal learning spaces. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 2021.

[14] Vishakh Padmakumar and He He. Does writing with language models reduce content diversity? *arXiv preprint arXiv:2309.05196*, 2023.

[15] Allison Parrish, Adam Simon, and Tim Szetela. Rewordable: The uniquely fragmented word game. `https://rewordable.com`, 2017. Accessed: 2023-10-28.

[16] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. Learning transferable visual models from natural language supervision. In *International Conference on Machine Learning*, pages 8748–8763. PMLR, 2021.

[17] Joel Simon and Ryan Murdock. New Words. `https://joelsimon.net/new-words.html`, 2023. Accessed: 2023-10-28.

[18] Thorny Games. Dialect: A game about language and how it dies. `https://thornygames.com/pages/dialect`, 2018. Accessed: 2023-10-28.

[19] Mike Treanor, Alexander Zook, Mirjam P Eladhari, Julian Togelius, Gillian Smith, Michael Cook, Tommy Thompson, Brian Magerko, John Levine, and Adam Smith. AI-based game design patterns. In *Proceedings of the International Conference on the Foundations of Digital Games*, 2015.

[20] Mike Vuolo. The meaning of Scrabble. https://www.slate.com/articles/podcasts/lexicon_valley/2012/03/lexicon_valley_the_role_of_language_in_scrabble_.html, 2012. Accessed: 2023-10-28.