Literature Review:
DAGGER: Data Augmentation for Generative Gaming in Enriched Realms

Noa JELSCH
CESI École d'Ingénieurs

June 2025

Summary

This literature review discusses the 2024 paper titled **DAGGER: Data Augmentation for Generative Gaming in Enriched Realms**, authored by Chris Callison-Burch, Ajay Patel, James Dennis, and Andrew Zhu from the University of Pennsylvania. The paper introduces a novel dataset and methodology aimed at enriching text-based game environments through large language model (LLM) augmentation. DAGGER extends the LIGHT dataset with GPT-4-generated characters, objects, and location-specific fiction, resulting in a corpus of over 259,000 words. It is designed to enhance generative and interpretive models in interactive storytelling.

Objectives and Motivation

The authors aim to:

- Address the lack of richly annotated datasets for training narrative-capable game agents.

- Enhance the connection between game state representation and narrative immersion.

- Provide a testbed for dual-directional modeling: generating stories from states and inferring states from fiction.

This work builds upon earlier imitation learning efforts by integrating narrative understanding into agent design, allowing for more human-like and believable game interactions.

Methodology

1. **Data Augmentation**: Using GPT-4 through the DataDreamer tool, the LIGHT environment is enriched with context-aware fictional descriptions, characters, and items.

2. **Dataset Design**: The final DAGGER corpus includes game state records, natural language narratives, and links between elements across rooms and scenes.

3. **Model Training**: Two transformer-based models are trained:

   - A forward generator to produce fiction from structured state data.
   - An inverse model to reconstruct game states from generated narratives.

4. **Evaluation**: Models are evaluated using perplexity and accuracy metrics to assess story coherence and state fidelity.

Results

- Enriched data improves narrative variety and consistency in forward generation.

- Inverse models trained on DAGGER can reconstruct complex states from narratives with notable accuracy.

- LLM-generated data offers scalable and controllable improvements over manual annotation, making DAGGER a useful training resource.

Discussion

DAGGER exemplifies how synthetic augmentation can support imitation and generative learning in game development. By coupling structured state representations with grounded fiction, it enhances both training data quality and narrative diversity. The dataset promotes future research into context-aware NPC behavior and interactive narrative engines.

Limitations include:

- Focus on text-only environments; no graphical or physical action grounding.

- Absence of player interaction studies or extrinsic task evaluation.

- Dependence on proprietary models (e.g., GPT-4) may limit reproducibility.

Conclusion and Future Work

DAGGER represents a significant advancement in narrative-rich game AI training. Future directions include:

- Adapting the framework for multimodal games.

- Incorporating human evaluations or user studies.

- Merging DAGGER's output with behavioural cloning pipelines for full-agent training.

Reference

Callison-Burch, C., Patel, A., Dennis, J., & Zhu, A. (2024). *DAGGER: Data Augmentation for Generative Gaming in Enriched Realms.* University of Pennsylvania. Retrieved from `https://www.cis.upenn.edu/~ccb/publications/dagger.pdf`