

УДК 004.8:004.912

**СИСТЕМА ИНТЕРАКТИВНОГО ПОВЕСТВОВАНИЯ С ПОДДЕРЖКОЙ  
КОНТЕКСТНО-ЗАВИСИМОГО СЦЕНАРИЯ: ТЕХНОЛОГИИ  
ГЕНЕРАЦИИ И АДАПТАЦИИ ТЕКСТА**

**ГУРЕВНИН Дмитрий Романович**

магистрант

*Научный руководитель:* **ЛЫСЕНКО Алексей Фёдорович**

кандидат технических наук, доцент

Донской государственной технической университет

г. Ростов-на-Дону, Россия

В статье рассматриваются подходы к разработке систем интерактивного повествования с использованием больших языковых моделей. Основное внимание уделяется вопросам организации контекста и поддержания согласованности сценария при взаимодействии с пользователем. Отдельно рассматриваются методы управления генерацией текста, а также возможности интеграции подобных решений с внешними цифровыми платформами. Показано, что для устойчивой работы интерактивных систем важно не только качество самой модели, но и наличие дополнительных механизмов контроля и обработки контекста.

**Ключевые слова:** интерактивное повествование, языковые модели, трансформеры, генерация текста, контекст.

В последние несколько лет интерес к интерактивным системам, способным формировать сюжет с учетом действий пользователя, заметно усилился. Подобные решения все чаще используются в игровой индустрии, образовательных проектах и в различных цифровых сервисах. Во многом это

стало возможным благодаря развитию больших языковых моделей, которые научились работать с текстом на более сложном уровне [9; 1; 6].

При этом подход к созданию сценариев постепенно меняется. Если раньше сюжет, как правило, задавался заранее и имел фиксированную структуру, то сейчас все чаще используется динамическая генерация. Это дает больше свободы, но одновременно приводит к новым трудностям. На практике нередко возникают ситуации, когда нарушается внутренняя логика повествования, «теряются» ранее заданные условия или поведение персонажей меняется без явных причин.

В данной работе рассматриваются современные подходы к созданию систем интерактивного повествования на основе языковых моделей. Основной акцент сделан на том, каким образом можно сохранить связность сценария при его динамическом формировании. В рамках исследования проанализированы архитектурные особенности моделей, методы работы с контекстом и способы контроля генерации.

Как показывают существующие исследования [1; 3; 6; 9], архитектура трансформеров хорошо подходит для генерации текста, однако вопросы долгосрочного управления контекстом и поддержания целостного повествования остаются до конца не решёнными.

## *2. Архитектурные основы генеративных моделей*

### *2.1. Особенности трансформеров*

Современные языковые модели основаны на архитектуре трансформера, в которой используется механизм внимания. Он позволяет учитывать зависимости между различными частями текста и формировать более связные ответы [6; 9].

Вместе с тем при работе с длинными сценариями возникают определённые сложности. В ряде случаев модель начинает «терять» ранее заданные условия или нарушает причинно-следственные связи между событиями. Это особенно заметно при длительном взаимодействии с пользователем.

### *2.2. Подходы к генерации текста*

Характер генерируемого текста во многом зависит от выбранного способа генерации. Детерминированные методы, дают более стабильный результат, но делают текст менее разнообразным. Вероятностные подходы, напротив, позволяют получать более вариативные ответы, но могут снижать их логическую согласованность [1; 2].

### *3. Управление контекстом*

#### *3.1. Внешние механизмы памяти*

Одной из ключевых задач при построении интерактивных систем является работа с контекстом. Поскольку сами языковые модели не обладают устойчивой долговременной памятью, на практике используются дополнительные механизмы хранения информации.

К таким механизмам относятся контекстные буферы, базы знаний и различные формы долговременной памяти [3; 10]. Например, в ряде систем применяется хранение описаний персонажей, событий и правил мира во внешней базе данных с последующим извлечением нужной информации при генерации ответа. Это позволяет уменьшить количество противоречий и повысить согласованность текста.

#### *3.2. Дообучение моделей*

Дополнительно используется дообучение моделей на специализированных корпусах. Это помогает лучше учитывать особенности конкретной предметной области, хотя в некоторых случаях может снижать универсальность модели.

#### *3.3. Использование RAG*

Отдельного внимания заслуживает подход Retrieval-Augmented Generation. Его суть заключается в том, что при генерации текста модель обращается к внешним источникам данных. На практике это может выглядеть как подгрузка информации о текущем состоянии сценария перед формированием ответа. Такой подход позволяет учитывать ранее произошедшие события и делает повествование более последовательным [3; 10].

### *4. Контроль генерации*

Управление процессом генерации обычно строится на сочетании нескольких подходов. К ним относятся задание стиля, введение ограничений и использование обратной связи от пользователя.

Как показывает практика, формализация правил мира и учет пользовательских оценок позволяют сделать поведение системы более устойчивым. При этом предсказуемость результата возрастает, что особенно важно для сценариев, где требуется соблюдение логики происходящего [1; 2; 5; 7].

Особенно заметна роль этих механизмов при интеграции систем с внешними платформами. В таких условиях любые несоответствия или ошибки в тексте становятся более очевидными, поэтому требования к стабильности работы модели возрастают.

#### *5. Этические аспекты*

Развитие интерактивных систем на основе языковых моделей сопровождается рядом этических вопросов. Это связано с высокой степенью автономности таких систем и использованием персонализации.

В отдельных случаях модель может воспроизводить нежелательные или неточные формулировки. При этом интерактивный формат делает такие ответы более значимыми для пользователя, поскольку они воспринимаются как часть осмысленного диалога. Из-за этого возрастает значение прозрачности работы системы. Пользователь должен понимать, каким образом формируются ответы и какие данные используются для персонализации. Также важно предоставить возможность управлять этими настройками, включая просмотр и удаление накопленной информации.

#### *6. Заключение*

Интерактивные системы, основанные на больших языковых моделях, уже сегодня находят применение в игровой индустрии, образовательных проектах и различных цифровых сервисах. При этом их практическое использование связано с рядом особенностей, которые необходимо учитывать на этапе

проектирования. Проведённый анализ показывает, что в основе подобных решений лежат трансформерные модели, обеспечивающие достаточно высокое качество генерации текста. Однако при работе с протяжёнными сценариями и сложными причинно-следственными связями они могут демонстрировать нестабильность, что требует использования дополнительных средств управления контекстом и построения более сложной архитектуры системы.

Как показывает практика, ключевую роль играет организация контекста. Использование внешней памяти, специализированных хранилищ знаний и подходов, основанных на Retrieval-Augmented Generation, позволяет компенсировать ограничения модели и расширить доступный объем информации. Это особенно важно при длительном взаимодействии с пользователем, когда необходимо сохранять сведения о событиях и персонажах, не допуская противоречий в повествовании.

Не менее важным является контроль процесса генерации. Выбор стратегии формирования текста, задание стилевых ограничений и учет пользовательской обратной связи позволяют в определенной степени управлять как вариативностью, так и предсказуемостью результатов. Это напрямую влияет на сохранение логики сценария и целостности образов персонажей [1; 2; 5; 7].

Дополнительные возможности открываются при интеграции интерактивных систем с внешними цифровыми платформами и сервисами. В таких условиях становится возможной более глубокая адаптация сценария под конкретного пользователя. Персонализация, основанная на анализе поведения, действительно повышает вовлеченность, однако одновременно усиливает требования к прозрачности работы системы и учету этических аспектов [7; 8; 4].

Таким образом, развитие интерактивных систем повествования связано не только с совершенствованием самих языковых моделей, но и с формированием комплексных архитектурных решений. Такие системы объединяют механизмы памяти, инструменты контроля генерации, средства интеграции с внешними

платформами и элементы персонализации, обеспечивая более устойчивое и согласованное взаимодействие с пользователем.

### **СПИСОК ЛИТЕРАТУРЫ**

1. Brown T.B. et al. Language Models are Few-Shot Learners // Advances in Neural Information Processing Systems. 2020. Vol. 33. P. 1877–1901.
2. Holtzman A. et al. The Curious Case of Neural Text Degeneration // ICLR. 2020.
3. Lewis P. et al. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks // Advances in Neural Information Processing Systems. 2020. Vol. 33. P. 9459–9474.
4. Mei F., Wang X. Interactive Story Generation Using Language Models // материалы конференции ACM.
5. Ouyang L. et al. Training Language Models with Human Feedback // NeurIPS (preprint arXiv, далее NeurIPS proceedings), 2022.
6. Raffel C. et al. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer // Journal of Machine Learning Research. 2020. Vol. 21. P. 1–67.
7. Roller S. et al. Recipes for Building an Open-Domain Chatbot // EACL. 2021.
8. Urbanek J. et al. Learning to Speak and Act in a Fantasy Text Adventure Game // EMNLP. 2019.
9. Vaswani A. et al. Attention Is All You Need // Advances in Neural Information Processing Systems. 2017. Vol. 30.
10. Yarats D. et al. Hierarchical Memory for Long-Context Language Models // arXiv preprint. 2023.

# MODERN INTERACTIVE STORYTELLING SYSTEMS WITH CONTEXT-AWARE SCENARIOS

**GUREVNIN Dmitry Romanovich**

Master's Student

*Research Supervisor:* **LYSENKO Alexey Fyodorovich**

Candidate of Technical Sciences, Associate Professor

Don State Technical University

Rostov-on-Don, Russia

This article examines approaches to developing interactive storytelling systems using large language models. The focus is on organizing context and maintaining narrative coherence during user interaction. The paper also discusses methods for controlling text generation, as well as the potential for integrating such solutions with external digital platforms. It is shown that for the stable operation of interactive systems, not only the quality of the model itself is important, but also the presence of additional mechanisms for controlling and processing context.

**Keywords:** interactive storytelling systems, large language models, transformers, context management.