

**Politechnika Wrocławskas
Wydział Informatyki i Telekomunikacji**

Kierunek: **Informatyka Stosowana**
Specjalność: **Projektowanie Systemów Informatycznych**

**PRACA DYPLOMOWA
MAGISTERSKA**

**Badanie wykorzystania sztucznej inteligencji
w procesie tworzenia dostosowującej się do
użytkownika narracji w grach komputerowych**

Kajetan Pynka

Opiekun pracy
Dr Maciej Walczyński

Słowa kluczowe: słowo1, słowo2, słowo3

STRESZCZENIE

 Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

ABSTRACT

 Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

SPIS TREŚCI

Wstęp	3
1. Narracja w grach	5
1.1. Historia narracji w grach komputerowych	5
1.1.1. Definicja narracji	5
1.1.2. Przedstawienie narracji w grach na przestrzeni lat	8
1.1.3. Prześledzenie rozwoju narracji na przykładzie serii Final Fantasy	15
1.2. Rodzaje narracji w grach komputerowych	19
1.2.1. Struktury narracyjne	19
1.2.2. Sposoby przedstawiania narracji	22
1.3. Systemy dialogowe w grach komputerowych	26
1.3.1. Popularne systemy dialogowe	26
1.3.2. Interaktywna fikcja - system poleceń	31
2. Sposoby generowania narracji	34
2.1. Wykorzystanie algorytmów sztucznej inteligencji do kreowania narracji	34
2.2. Wykorzystanie dużych modeli językowych (LLM) do kreowania narracji	38
3. Zaangażowanie gracza	43
3.1. Definicje i rodzaje zaangażowania	43
3.2. Sposoby pomiaru zaangażowania gracza	45
4. Planowany eksperyment	48
4.1. Projekt gry wykorzystanej w eksperymencie	48
4.2. Opis generatywnych agentów	51
4.3. Zaplanowany przebieg eksperimentu	54
5. Wyniki	55
Podsumowanie	56
Spis rysunków	57
Spis listingów	59
Spis tabel	60
Dodatki	61
A. Słownik	62
B. Instrukcje inworld.ai dla postaci	63

WSTĘP

Niniejszy szablon jest adaptacją szablonu opracowanego przez Pana Wojciecha Myszkę (patrz <https://kmim.wm.pwr.edu.pl/myszka/projekty>).

Został przetestowany w następujących narzędziach:

- Overleaf.com – wersja on-line; nie jest wymagana instalacja
- TeXStudio/TeXLive oraz TeXStudio/MiKTeX oraz TeXWorks/MiKTeX

Użycie pełnej wersji może wymagać (np. dla zestawu TeXWorks/MikTeX):

- instalacji Python 2.7+ wraz z pakietem Pygments; ścieżki do obu narzędzi powinny być ustawione w zmiennej środowiskowej PATH – pakiet jest używany do kolorowania słów kluczowych w listingach języków programowania
- w środowisku Latex należy włączyć opcję -shell-escape – jest wymagana dla pakietu “minted” Latexa

ZAKRES PRACY

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

CEL PRACY

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

STRUKTURA PRACY

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

1. NARRACJA W GRACH

Niniejszy rozdział ma na celu dokonanie przeglądu gier komputerowych na przestrzeni lat, ze szczególnym naciskiem na ewolucję sposobów oraz form narracji przedstawianych w tych grach. Wyszczególnione zostaną również naistotniejsze struktury i rodzaje narracji, które są współcześnie wykorzystywane. Dodatkowo, nastąpi krótki przegląd najpopularniejszych technik prezentacji narracji. Na koniec nakreślone zostaną systemy dialogowe wykorzystywane przez gry komputerowe.

1.1. HISTORIA NARRACJI W GRACH KOMPUTEROWYCH

Aby zrozumieć istotę narracji w grach komputerowych, należy przede wszystkim określić co może kryć się pod tym pojęciem. Pozwoli to dokonać przeglądu wybranych tytułów i wyciągnąć z tego przeglądu wnioski. Żeby udowodnić rozwój w sposób prezentowania narracji na przestrzeni lat, prześledzone zostały części jednej z serii gier — "Final Fantasy" — wydawanej od roku 1987.

1.1.1. Definicja narracji

Pojęcie narracji i samo jej występowanie w grach komputerowych jest kwestią sporną w literaturze od lat. Barry Ip, w swojej pracy [15], dokonuje wyróżnienia trzech słów ściśle powiązanych ze sobą: *historia*, *fabuła* oraz *narracja*. Na potrzeby jego badań historia zdefiniowana została następująco:

... *sekwencja zdarzeń obejmujących byty*. [15]

Związana z historią jest również fabuła, która została określona przez Arystotelesa jako:

... *organizacja zdarzeń*. [15]

Sama narracja, ściśle powiązana z dwoma poprzednimi terminami, wyrażona została w sposób następujący:

... *reprezentacja zdarzenia lub serii zdarzeń*. [15]

W ramach tej pracy, można przyjąć wszystkie te pojęcia jako istotne i na tyle bliskie siebie, że mogą być wykorzystywane zamiennie.

Jakub Majewski sugeruje, że debatowanie nad istnieniem narracji jest odpowiednie dla niektórych gier, a dla niektórych nie [20]. Rozdzielenie bowiem tych form przekazu, które można zaliczyć do treści fabularnej, nie jest takie oczywiste. Przytoczyć można przykład *Space Invaders* (1977) — gra nie przytacza żadnego opisu w formie tekstu, skupiając się wyłącznie na rozgrywce. Na podstawie samego tytułu można jednak przypuścić, że dokonuje się pewnego rodzaju inwazja, a stoją za nią przybysze z kosmosu [20].

Ten przykład pokazuje, że granica między grami posiadającymi narrację a tymi, w których jest ona nieobecna, może być płynna. Nawet gry pozbawione bezpośrednich opisów fabularnych mogą zawierać pewne nawiązania narracyjne, które wynikają z innych elementów, takich jak tytuł czy grafika. W związku z tym, podział na gry z narracją i bez narracji może być problematyczny, ponieważ elementy narracyjne mogą przejawiać się w różnych formach i stopniach w różnych grach. Jako że nie jest to główny problem poruszany w niniejszej pracy to wszystko co może być elementem narracyjnym, jest za taki uznawany.

Do budowania narracji w grach wykorzystane mogą być wzorce znane z literatury. Przykładem takiego wzorca jest "*Podróż bohatera*"[15], który opisuje 12 kluczowych etapów, odgrywających istotną rolę w budowie angażujących historii (Tabela 1.1). Blisko powiązana z "*Podróżą bohatera*" jest znana struktura trzech aktów opisana przez Arysotelesa, która zakłada podział utworu na początek, środek i koniec[15]. Jest to bardzo elastyczna a zarazem bardzo ogólna metoda podziału. Zasadniczo w każdym utworze dałoby się bowiem w pewien sposób wyodrębnić te akty.

Struktury te pozwalają projektantom fabuły konstruować spójny świat fikcji — niezależnie od formy w jakiej zostanie zaprezentowana odbiorcom. Takowa może być adaptowana zarówno do powieści, jak i do materiału filmowego czy też gier komputerowych.

Etap	Opis
1. Zwyczajny świat	Gracz po raz pierwszy spotyka bohatera i zapoznaje się z jego pochodzeniem, zazwyczaj za pośrednictwem historii drugoplanowej
2. Wezwanie do przygody	Wskazówka, że bohater opuści zwykły świat, by rozpocząć nową przygodę. Ten etap działa jak katalizator, który uruchamia główny wątek fabularny
3. Odrzucenie wezwania	W tradycyjnej strukturze monomitu bohater odrzuca początkową propozycję opuszczenia zwykłego świata i rozpoczęcia misji, zwykle w chwili wątpliwości lub niepewności
4. Spotkanie z mentorem	Gdy bohater decyduje się na podjęcie zadania, mentor dostarcza mu informacji potrzebnych do podjęcia decyzji. Mentorem może być wszystko, co dostarcza informacji - brodaty starzec, robot, biblioteka, doświadczenia z przeszłości i tak dalej
5. Przekroczenie pierwszego progu	Bohater przechodzi z bezpiecznego zwykłego świata do nowego, niebezpiecznego i nieznanego świata poszukiwań
6. Testy, sprzymierzeńcy i wrogowie	Faza ta jest zwykle największą częścią fabuły gry, ponieważ gracz poznaje wszystkie główne postacie
7. Podejście do najgłębszej jaskini	Jest to miejsce, w którym bohater znajduje nagrodę, której szuka - taką jak zdobycie niezbędnej umiejętności, broni lub opanowanie wszystkiego, co napotkał do tej pory. Zazwyczaj ma to miejsce pod koniec gry. Głównym celem tej części historii jest przygotowanie bohatera do ostatecznej bitwy
8. Próba	To tutaj bohater staje do ostatecznej walki ze swoim nemezis lub "ostatecznym bossem". Nemezis może pojawić się jako byt fizyczny (osoba lub przedmiot) lub niefizyczny (czas, intensywność lub trudność)
9. Nagroda	Wiele gier kończy się w tym momencie, gdy wróg zostaje pokonany, a nagrodą jest zazwyczaj końcowa cut-scenka opisująca, co dzieje się z bohaterem po jego triumfie
10. Droga powrotna	Niektóre gry pozwolą graczowi powrócić do zwykłego świata po otrzymaniu nagrody, ale może nie być możliwe, aby bohater z powodzeniem zintegrował się ze starym światem
11. Wkszeszenie	Ta część historii odpowiada na wszelkie pytania bez odpowiedzi, takie jak konsekwencje misji, potencjalne konflikty, które mogą pojawić się w przyszłych sequelach, lub wszelkie testy, którym bohater musi stawić czoła przed końcem. Może mieć również formę ostatecznego zwrotu akcji, jako coś nieoczekiwane przez widzów
12. Powrót z nagrodą	Jest to ostatni etap historii, w którym bohater w końcu powraca do zwykłego świata i widzi korzyści płynące z jej nagrody. Bohater może porównać swoje życie przed i po wyprawie, aby zobaczyć, jak wszystko się zmieniło

Tabela 1.1: Dwanaście etapów wzorca narracyjnego "Podróży bohatera" [15]

1.1.2. Przedstawienie narracji w grach na przestrzeni lat

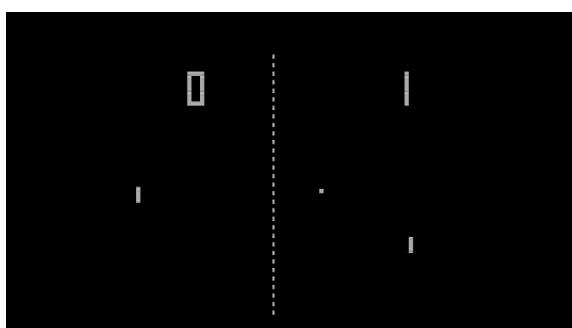
Kamienie milowe w początkach branży gier wideo to: Spacewar (Rys 1.1) - pierwsza interaktywna gra z 1962 roku, Magnavox Odyssey (1972) - pierwszy domowy system gier podłączany do telewizora, a także Pong (Rys 1.2) od Atari (1972) i przenośne gry LED Mattela (1977)[30].

Gra "Spacewar" (Rys 1.1) przedstawia dwa statki kosmiczne, które w obrębie studni potencjału grawitacyjnego ("gravity well") prowadzą ze sobą starcie. Jeden ze statków nazywany jest "igłą" a drugi "klinem". Oba są sterowane przez graczy, którzy mają do dostępu ograniczoną amunicję i paliwo do nawigowania. Cała rozgrywka prowadzona jest na planszy 2-wymiarowej, gdzie tło stanowią gwiazdy. Gra nie posiada żadnej formy narracji, natomiast była istotnym elementem dalszego rozwoju branży.



Rys. 1.1: Spacewar (1962)

W ramach rozgrywki w "Pong" (Rys 1.2) mamy do czynienia z symulatorem tenisa stołowego. Dwójka graczy steruje paletkami poruszającymi się pionowo. Za pomocą tych paletek odbijają piłkę na stronę przeciwnika. Jeśli ten nie odbije jej z powrotem, to uderzający zdobywa punkt. Wygrywa pierwszy gracz, który uzyska 11 punktów. Podobnie jak w przypadku "Spacewar", "Pong" nastawiony jest na rozgrywkę dwuosobową i nie posiada żadnej formy narracji.



Rys. 1.2: Pong (1972)

Lata 70. przyniosły rozwój firm jak Atari, Nintendo i Sega oraz pierwsze hity salonów gier np. Pacman (1980), który sprzedał 300 000 sztuk na całym świecie[30].

W "Breakout" (Rys 1.3) gracz steruje paletką poruszającą się poziomo i stara się zniszczyć położoną wyżej ścianę z cegiełek. Ściana składa się z ośmiu rzędów kolorowych bloczków. Używając pojedynczej piłki należy zbić jak najwięcej cegiełek (przy kontakcie piłki z cegielką zostaje ona zniszczona). Grający posiada trzy życia i w ramach nich musi wyczyścić dwie ściany. Gracz traci życie jeśli nie odbije piłki wracającej do niego. Rozgrywka ta została zaplanowana na maszyny *arcade* z myślą o zdobywaniu jak najwięcej punktów. Nie da się dostrzec w jej przypadku żadnej formy fabuły.



Rys. 1.3: Breakout (1976)

"Space Invaders" (Rys 1.4) to gra akcji opracowana i wydana w Japonii przez Taito. Gracz steruje działem laserowym umieszczonym na dole ekranu, które porusza się poziomo. Kosmici ułożeni w 5 rzędów po 11 obiektów przemieszczają się grupowo w lewo i prawo, schodząc niżej gdy dotkną krawędzi ekranu. Celem gry jest zestrzelenie wszystkich kosmitów przez gracza, posiadającego trzy życia. Obcy wystrzelują swoje pociski, które przy trafieniu w gracza zabierają mu jedno życie. Gra kończy się natychmiastowo w momencie gdy najeźdźcy dotrą do dołu ekranu. Tak jak w przypadku "Breakout" mamy do czynienia z rozgrywką nastawioną na maszyny *arcade*, a co za tym idzie na zdobywanie punktów. Oprócz kwestii poruszanych w podsekcji 1.1.1, nie występują inne przesłanki fabularne.



Rys. 1.4: Space Invaders (1978)

W latach 80. nastąpił boom konsol domowych - Nintendo NES, Sega Master System, Atari 7800[30]. Obecnie kiedy większość graczy myśli o grach *retro* to ma na myśli między innymi właśnie tytuły wyprodukowane na te serie konsol.

Jedną z najbardziej znanych w popkulturze gier jest "Super Mario Bros." (Rys 1.5). Gracz wciela się w rolę tytułowego Mario (w wersji jednoosobowej), który ma jako główne zadanie obrane ocalenie księżniczki. W tym celu pokonuje kolejne krainy (poziomy) oraz przeciwników. Gra składa się z ośmiu światów, gdzie każdy z nich jest dodatkowo podzielony na cztery poziomy. Mamy więc do czynienia z wielopoziomową formą rozrywki, gdzie światy różnią się między sobą ze względu na warstwę wizualną, dźwiękową jak i ze względu na występujących przeciwników czy przeszkody. Są to pewnego rodzaju zalążki narracji (która została wykreowana przez świat). Jedyną formą pisemnej fabuły jest tekst występujący po ukończeniu poziomu (Rys 1.5b).

(a) Ekran tytułowy



(b) Ukończenie poziomu



Rys. 1.5: Super Mario Bros. (1985)

"The Legend of Zelda" (Rys 1.6) to gra przygodowa, w której główną postacią sterowaną przez gracza jest Link. Jego zadaniem jest zebranie ośmiu fragmentów Trójkątnej wiedzy (ang. *Triforce of Wisdom*) by uratować księżniczkę Zeldę. Przy rozpoczęciu rozgrywki

graczowi przedstawiany jest ekran ze wstępem fabularnym (Rys 1.6a). Gra posiadała również dedykowaną instrukcję, która na zasadzie poradnika podawała wskazówki dotyczące rozgrywki. Tak jak w przypadku "Super Mario Bros.", występuje podział na poziomy, a co za tym idzie zmienia się oprawa audio-wizualna jak i spotykani przeciwnicy. W trakcie rozgrywki możemy napotkać na postacie NPC (ang. *non-playable character*), które komunikują się za pomocą krótkiego stwierdzenia (Rys 1.6b). Widoczne są również załączki motywów "otwartego świata", gdzie gracz zwiedza świat i jego elementy w dowolnej kolejności.



Rys. 1.6: The Legend of Zelda (1986)

Dekadę później gry komputerowe PC zyskały popularność dzięki tytułom jak Doom, a na rynku pojawiły się PlayStation i Nintendo 64. Koniec XX wieku to także rozwój przenośnych gier na fali sukcesu serii Pokemon.

Pierwszym tytułem opisany w tej podsekcji, który operuje na perspektywie 3-osobowej jest "Crash Bandicoot" (Rys 1.7). W wydanej na platformę PlayStation w 1996 roku grze gracz steruje tytułowym Crash'em Bandicoot'em. Rozgrywkę otwiera tzw. cut-scenka (wyjaśniona również w podsekcji 1.2.2), czyli krótki film wprowadzający do fabuły. To właśnie dzięki niej gracz może usłyszeć wypowiadane imię głównego przeciwnika Crash'a, doktora Neo Cortex'a. Szalony naukowiec prowadził badania na lokalnej faunie i pragnął by Crash stał się generałem jego armii. Crash zostaje jednak odrzucony przez maszynę do prania mózgu i ucieka z zamku Cortex'a. Pod koniec filmu przedstawiona zostaje też "kobieta Bandicoot", na której mają zostać przeprowadzone dalsze testy. W ten sposób gracz odkrywa dlaczego Crash przemierza kolejne poziomy (by ocalić panią Bandicoot), dlaczego przeciwnicy są tacy a nie inni (przez testy doktora Neo Cortex'a) i co za tym idzie kto jest jego głównym rywalem. Dodatkowo, w ramach rozgrywki mamy do czynienia z perspektywą trójwymiarową, która jednak ulega zmianie pomiędzy poziomami. W grze

występują również dwa zakończenia, jedno wymaga perfekcyjnego przejścia gry, natomiast oba przedstawione są w formie cut-scenki.

(a) Cut-scenka otwierająca



(b) Crash i jego towarzysz Aku Aku



Rys. 1.7: Crash Bandicoot (1996)

Uznawaną za jedną z najlepszych czy też najbardziej kultowych gier komputerowych jest "Half-Life" (Rys 1.8) wydany w 1998 roku. Wprowadzająca do rozgrywki sekwenca przedstawia najważniejsze informacje, nie zabierając przy tym kontroli (gracz może przemieszczać się i rozglądać po kabinie pociągu). Nakreślona zostaje postać Gordon'a Freemana, 27-letniego doktora fizyki teoretycznej, który pracuje w placówce Black Mesa, położonej w Nowym Meksyku. Niemal cała warstwa fabularna zostaje wypowiedziana przez aktorów głosowych, bez wspierających napisów. Gracz w trakcie rozgrywki spotyka różnych NPC - i to właśnie one przekazują mu istotne informacje (posiadają również odpowiednie animacja poruszania ustami przy wypowiadaniu kwestii). Przejścia pomiędzy poziomami są maskowane przez m.in. podróże windą czy bardzo krótkie doczytywanie kolejnych fragmentów świata — co sprawia, że rozgrywka pozostaje bez przerwy imersyjna.

(a) NPC w monologu



(b) Napisy w sekwencji otwierającej



Rys. 1.8: Half-life (1998)

Nowe millennium przyniosło dalszy rozwój branży do rozmiarów dzisiejszej potęgi, poprzez stale pojawiające się innowacje sprzętowe i nowe przełomowe tytuły na różne platformy.

Przykładem idealnie obrazującym rozwój w podejściu do narracji jest gra "Life is Strange" (Rys 1.9) wydana w 2015 roku. Jest to gra przygodowa o charakterze fikcji interkatywnej (Więcej na ten temat w podsekcji 1.3.2), która wydawana była w formie epizodycznej. Każdy epizod przypominać może podział na odcinki znany z seriali telewizyjnych czy też akty w literaturze. Centralną postacią jest Maxine Caulfield, która odkrywa że potrafi cofnąć czas. Całość fabuły opiera się na motywie *efektu motyla*, gdzie podejmowane decyzje mogą rzutować na wydarzenia w przeszłości czy też na relacje pomiędzy postaciami. Z tego powodu narrację można zakwalifikować jako rozgałęzającą się (Patrz 1.2.1) i zarazem stanowi ona kluczowy element rozgrywki a nie tylko "imersyjny dodatek" do niej. System dialogów zaprezentowany w "Life is Strange" oparty jest na popularnej strukturze kołowej (Patrz 1.3.1) z motywem drzewa, tzn. konkretne wybory dialogowe wiążą się z gamą zupełnie nowych wyborów w dalszej części rozmowy (mechanika cofania czasu pozwala graczu rozgrywać dialog na inny sposób).

(a) Wybór z ostrzeżeniem o konsekwencjach



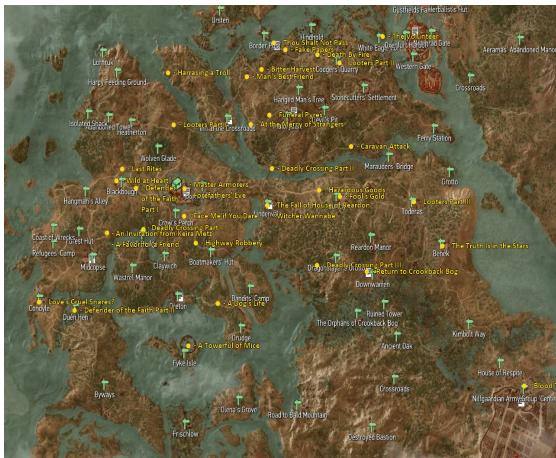
(b) System dialogowy kołowy



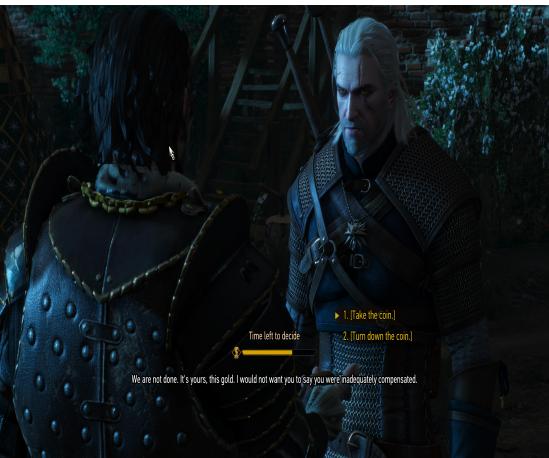
Rys. 1.9: Life is Strange (2015)

Jako wzór gry z otwartym światem niewątpliwie można uznać tytuł "Wiedźmin 3" (Rys 1.10) wydane w 2015 roku i wyprodukowane przez polskie studio CD Projekt Red. Akcja rozgrywka się w świecie opartym na powieściach i opowiadaniach Andrzeja Sapkowskiego a gracz wciela się w rolę Geralta z Rivii. Forma otwartego świata oznacza w tym przypadku tyle, że grający ma możliwość swobodnego poruszania się po krainach i jest też w stanie podejmować się różnych zadań w "niemal" dowolnej kolejności. Sprawia to, że gracz nie czuje się niczym aktor odgrywający kolejne sceny a za to jest reżyserem własnych przygód. System dialogowy zorganizowany jest w formę menu wyborów, przy czym na niektóre zdarzenia gracz ma ograniczony czas odpowiedzi (Przedstawione na Rys. 1.10b). Dodatkowo, niektóre opcje mogą wymagać posiadania określonej ilości pieniędzy czy też odblokowania określonych umiejętności.

(a) Mapa świata (żółte punkty - zadania)



(b) Dialog z czasem na decyzję



Rys. 1.10: Wiedźmin 3 (2015)

1.1.3. Prześledzenie rozwoju narracji na przykładzie serii Final Fantasy

Seria "Final Fantasy" zadebiutowała w 1987 roku na konsoli NES, a samym twórcą gry był Hironobu Sakaguchi. Na przestrzeni kolejnych lat powstało 14 numerowanych odsłon serii oraz wiele spin-offów [17]. W głównej serii każda gra nie miała nic wspólnego z poprzednią pod względem fabuły, postaci czy uniwersum. Pojawiały się wspólne motywy i stworzenia, ale na ogół każde "Final Fantasy" stanowiła odrębną, zamkniętą przygodę[17].

Seria przeżywała i nadal przeżywa swoistego rodzaju ewolucję narracyjną. Elementy, które stanowiły główną część fabularną, zeszły na dalszy plan. Początkowo mało istotni bezimienni protagonisti zostali zastąpieni przez postacie angażujące się w dialogi i rozwijające się na przestrzeni gry [17].

Aby prześledzić proces zmian w serii, dokonany zostanie przegląd kolejnych odsłon "Final Fantasy" na podstawie informacji zebranych przez Kevin'a Kryah[17] oraz Hayes'a Madsen'a[18].

W pierwszej odsłonie (Rys. 1.11) fabuła była dość prosta - czterech sterowalnych bohaterów, znanych jako *Wojownicy Świata*, przybywają do królestwa *Cornelia*, niosąc mistyczne kule. Zostają poinformowani, że muszą pokonać cztery żywioły, by przywrócić światu równowagę [17]. W samej grze *Wojownicy Świata* są niemymi postaciami, natomiast występują NPC (ang. *non-playable character*), którzy w formie monologów przekazują graczy informacje. Monologi te przedstawione są jedynie w formie tekstuowej, nie występuje bowiem zjawisko aktorstwa głosowego (po ang. *voice acting*). Cut-scenki (patrz 1.2.2) są bardzo prymitywne, na zasadzie poruszania kamerą i wykorzystaniu animacji występujących w grze (nie są prerenderowane). W ramach ekranu startowego gry przedstawione zostaje graczy krótkie wprowadzenie fabularne (Rys. 1.11a).



Rys. 1.11: Final Fantasy I (1987)

Żadne z kolejnych "Final Fantasy" nie wniosły zbyt wiele pod względem złożoności fabuły. "Final Fantasy IV" wprowadziła większy nacisk na rozwój postaci poprzez cut-scenki i wątki bohaterów. Głównym elementem nadal pozostawała walka z "ostatecznym złowrogim bossem", lecz przy okazji gracz był w stanie poznać główne postacie[17].

"Final Fantasy VI" poszła o krok dalej - rolę postaci zaczęła przesłaniać nadziedną fabułę. Sceny takie jak próba samobójcza Celes czy rzut monetą między Edgarem i Sabinem niewiele wniosły do głównego wątku, ale dobrze charakteryzowały bohaterów w sposób, którego nie sposób opisać samymi dialogami. [17].

W ramach "Final Fantasy VII" (Rys. 1.12) doszło do przełomu graficznego, ponieważ rozgrywka odbywała się w świecie trójwymiarowym. Dodatkowo, ta część okraszona została pełnoprawną cut-scenką otwierającą.

"Final Fantasy VII" i VIII kontynuowały trend nastawienia na postaci. Ich losy stały się głównym tematem, a walka dobra ze złem została odstawiona na dalszy plan. "Final Fantasy VIII" wydawała się wręcz niezainteresowana głównym wątkiem, a jej zakończenie koncentrowało się na zamknięciu wątków postaci i zostało zwieńczone pocałunkiem na balkonie [17].

W swej istocie, "Final Fantasy VIII" jest formą opowieści miłosnej. Mimo występowania oczywistych elementów fantastycznych czy fikcyjnych, to głównym motywem przyciągającym uwagę gracza jest miłość Squalla i Rinoi[18].

Hayes Madsen sugeruje, iż to właśnie przez występowanie takich osobistych historii czy relacji między postaciami fabuła "Final Fantasy VIII" jest tak często wspominana. Twórcy przedstawili bowiem występujące pomiędzy głównymi postaciami poczucie koleżeństwa i ich wspólnego rozwoju[18].

(a) Klatka cut-scenki wprowadzającej



(b) Dialogi pomiędzy postaciami



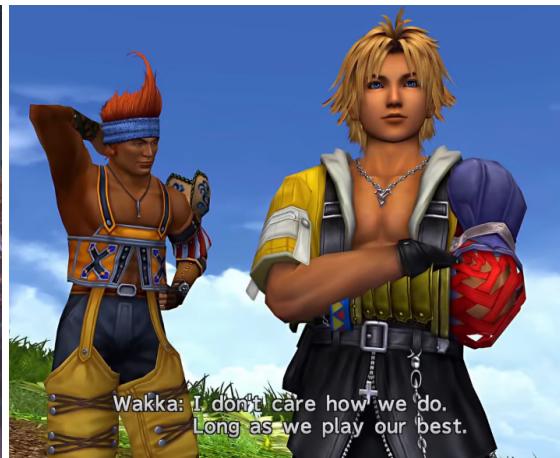
Rys. 1.12: Final Fantasy VII (1997)

Gdy seria wkroczyła w XXI wiek, gry oddalały się jeszcze bardziej od swych korzeni. "Final Fantasy IX" kontynuowała nastawienie na postacie, dodając klimat europejskich baśni czy aluzji do Szekspira i Carrola. "Final Fantasy X" (Rys. 1.13) podejmowała tematy religii i człowieczeństwa, prezentując jednocześnie jedne z najbardziej artystycznych wizji w serii[17]. Rozgrywka została urozmaicona bardzo wieloma zaawansowanymi przerywnikami filmowymi. Postacie wydawać się mogą zdecydowanie bardziej realistyczne czy "żywsze" ze względu na udział aktorów głosowych, począwszy od tej części.

(a) Cut-scenka z gry

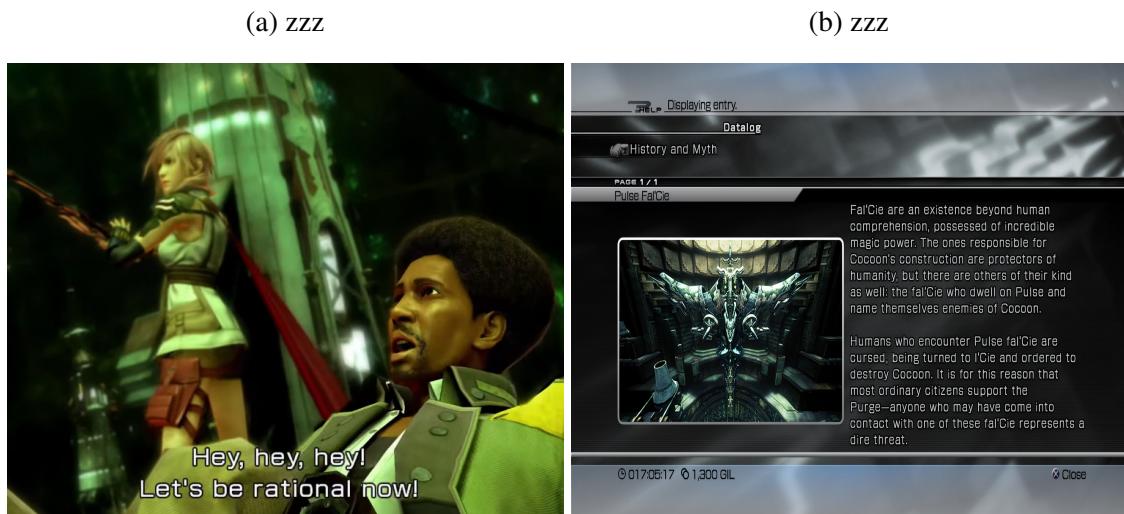


(b) Dialogi pomiędzy postaciami



Rys. 1.13: Final Fantasy X (2001)

"Final Fantasy XIII" (Rys. 1.14) w pełni skupiło się na rozwoju postaci i formie wizualnej, natomiast sensowność fabuły pozostawiała wiele do życzenia. W rzeczywistości większość istotnych informacji nie jest dostarczana w dialogach, ale za pośrednictwem wpisów do dziennika, znajdywanych przez gracza w trakcie rozgrywki. To raczej postacie zajęły centralne miejsce, a bezsensowna fabuła starała się jedynie związać ze sobą wszystkie wątki bohaterów[17].



Rys. 1.14: Final Fantasy XIII (2009)

Seria "Final Fantasy" rozwija się po dzień dzisiejszy, a twórcy nadal próbują używać różnych rozwiązań by wyróżnić kolejne tytuły. Przykładowo, w najnowszej odsłonie serii — "Final Fantasy XV" — deweloperzy odeszli od konwencji JRPG (ang. *Japanese role-playing game*) na rzecz gry akcji z otwartym światem.

Na przestrzeni lat seria "Final Fantasy" przeszła znaczącą ewolucję pod względem roli i znaczenia narracji. Niezależnie od zmian w strukturze rozgrywki czy sposobie prezentowania fabuły, widoczny jest stale rosnący wpływ elementów narracyjnych na całokształt kolejnych odsłon serii.

1.2. RODZAJE NARRACJI W GRACH KOMPUTEROWYCH

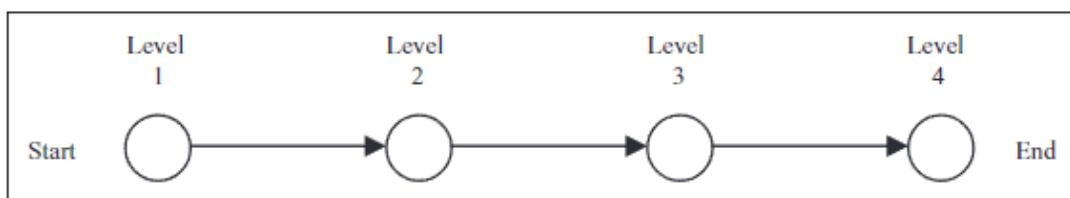
Poniższa sekcja ma na celu uporządkowanie — znanych z literatury czy też istniejących przykładów — struktur narracyjnych, za pomocą których opisać można sekwencję kolejnych wydarzeń w grze. Dodatkowo, przedstawione zostaną kluczowe sposoby czy też techniki, za pomocą których twórcy budują wirtualne światy fabularne.

1.2.1. Struktury narracyjne

Pod pojęciem *struktury narracyjnej* rozumiane jest **uporządkowanie** wydarzeń odbywających się w grze, które niosą jakiekolwiek przesłanki fabularne. Nie oznacza to, że każde wydarzenie musi się odbyć — może być to bowiem zależne od decyzji podjętych przez graczącego. Wyszczególnione zostaną trzy klasyczne struktury, które zapewniają dość płynny przebieg historii, a są to: *liniowa*, *łańcuch pereł*, *rozgałęziającą się*[30][20][15]. Z zakresu mniej oczywistych architektur dodatkowo warto wspomnieć o modelach *parku rozrywki* i *cegiełek*[20].

Liniowa

Jest to forma przekazu znana bardzo dobrze z literatury czy też kinematografii. Srowadza się ona bowiem do jednego ciągu zdarzeń, gdzie odbiorca nie ma wpływu na dalszy przebieg fabuły lecz jest on raczej pasywnym obserwatorem odgrywających się scen. W przypadku książki czy filmu jest to naturalne podejście ze względu na brak interaktywności. Jeśli chodzi o gry komputerowe, to strukturę tą można zaobserwować zwłaszcza w "starszych" tytułach (Np. wspomniany wcześniej "Crash Bandicoot" - patrz 1.1.2). Gra zasadniczo może być ukończona na jeden sposób — tak jak to zaplanowali projektanci[30] (Patrz Rys. 1.15).

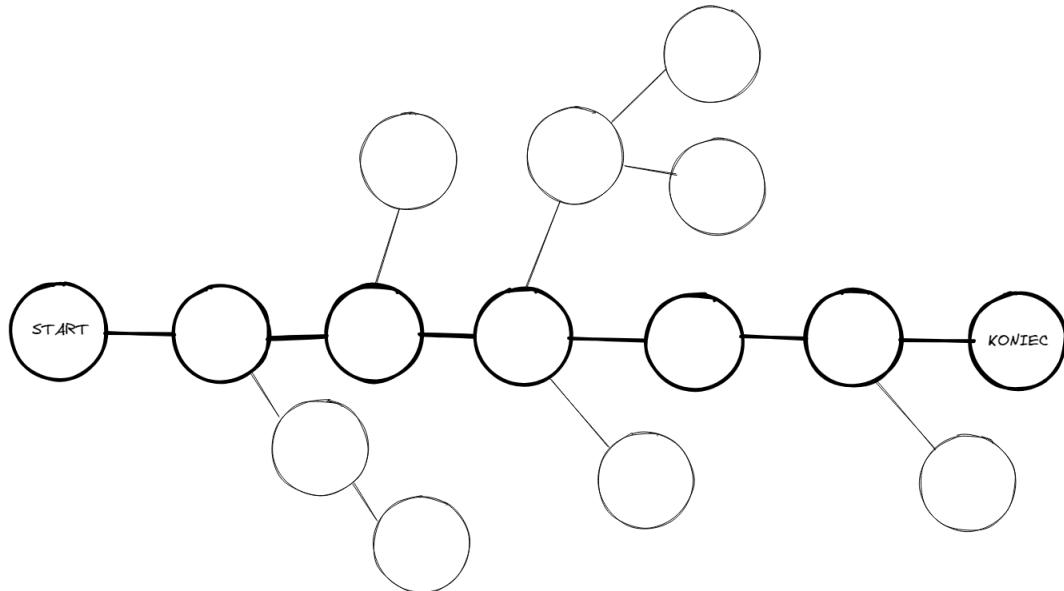


Rys. 1.15: Liniowa struktura gry[15]

Łańcuch pereł

W ramach tego modelu gracze uzyskują pewnego rodzaju "*iluzję*" wyboru. Występują podczas rozgrywki momenty swobody, gdzie grający mają poczucie wpływu na dalszy przebieg fabularny. W rzeczywistości jednak podejmowane przez nich decyzje mogą nie posuwać historii na przód, a same postępy narracyjne nadal znajdują się pod kontrolą

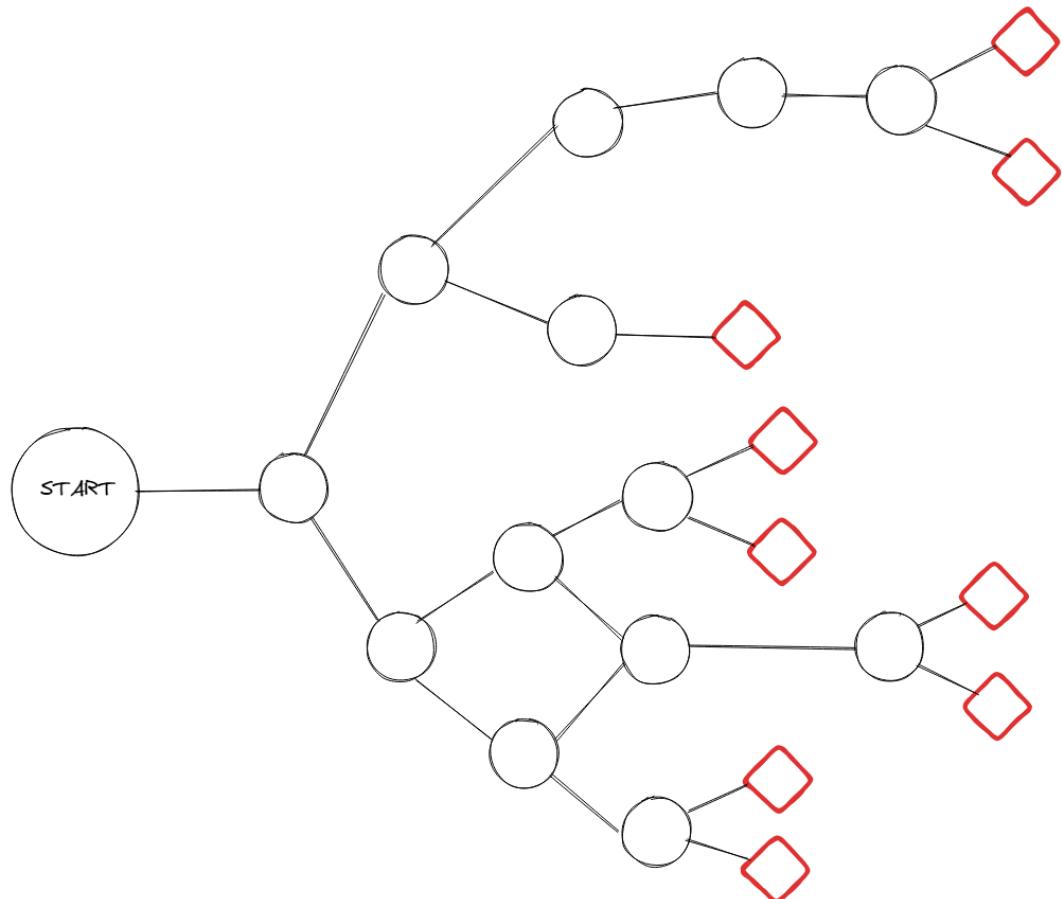
projektantów gry[20]. Tak jak przedstawiono na rysunku 1.16 — mogą występować tymczasowe rozgałęzienia wychodzące od głównej sekwencji fabularnej, natomiast zostają one w końcu urwane, a gracz zobowiązany jest do kontynuowania przygody według zaplanowanej historii. Jako przykład może służyć "The Legend of Zelda" (Patrz 1.1.2), gdzie grający może zwiedzać dane poziomy w dość dowolny sposób, natomiast musi ostatecznie trafić na główną ścieżkę by móc dokonać postępu.



Rys. 1.16: Struktura łańcucha pereł

Rozgałęziająca się

Metodą, która oferuje graczowi istotny wpływ na przebieg dalszej rozgrywki jest zdecydowanie model rozgałęziający się. Jak sama nazwa wskazuje, historia nie trzyma się jednej konkretnej wersji lecz jest w stanie "*rozgałęziać się*" w wielu innych możliwych kierunkach. Może to wynikać z jawnych decyzji podejmowanych przez gracza w istotnych momentach lub też ze względu na sposób w jaki podchodzi on do rozgrywki (np. może pomijać pewne elementy świata)[20]. W wyniku takich rozgałęzień wytwarza się pewnego rodzaju "sieć możliwości fabularnych"[20] — które z reguły muszą być uprzednio przygotowane przez projektantów gry. Struktura ta została przedstawiona na rysunku 1.17 — akcja zaczyna się w jednym punkcie, potem w wyniku decyzji istniejących w grze występują rozgałęzienia, które ostatecznie prowadzą do potencjalnie różnych zakończeń (oznaczonych czerwonymi rombami). Przykładem realizującym ten model może być wspomniany wcześniej tytuł "Life is Strange" (Patrz 1.1.2).



Rys. 1.17: Struktura rozgałęziająca się

Park rozrywki

Struktura ta z założenia bardzo przypomina model rozgałęziający się, natomiast w tym przypadku możemy mówić o narracji rozwijającej się ze względu na przestrzeń a nie czas[20]. Przykładowo, zwiedzając świat gry gracz może napotkać postać NPC, która otworzy przed nim nową gałąź fabularną (np. poprzez zlecenia zadania do wykonania)[30]. Model ten jest bardzo popularny dla gier z otwartym światem, przykładem może być "Wiedźmin 3" (Patrz 1.1.2).

Cegiełki

W ramach niektórych tytułów twórcy nie skupiają się na stworzeniu narracji możliwej do doświadczenia przez graczącego, lecz na pewnym systemie części, za pomocą których gracz sam jest w stanie tworzyć historię. Części te nazywane "*cegielkami*" (ang. *building blocks*)[20] są wykorzystywane przez graczącego do tworzenia własnej narracji. Przykładem tego rodzaju rozgrywki może być tytuł "The Sims" (2000), w którym to gracz tworzy i steruje rodziną — a co za tym idzie, kieruje ich historią życia.

1.2.2. Sposoby przedstawiania narracji

Oprócz zaplanowania i rozłożenia fabuły gry na części — przy użyciu kombinacji struktur opisanych w poprzedniej sekcji — istotną kwestią pozostaje wybór w jaki sposób dane sekwencje fabularne mają zostać przekazane graczowi. W ramach tej sekcji opisane zostaną najważniejsze techniki prezentowania narracji.

Cut scenki

Jednym z najbardziej widowiskowych sposobów prezentowania treści fabularnej jest zdecydowanie cut scenka, która zdefiniowana została przez Glassnera (2004) następująco:

... pre-renderowany fragment wideo ... czasami renderowany w czasie rzeczywistym przy użyciu sprzętu komputerowego lub konsoli. Podczas odtwarzania możliwość interakcji gracza zostaje zawieszona, a on sam staje się biernym widzem na czas trwania sceny[15]

Cut scenka jest zatem po prostu fragmentem wideo. Można dokonać pewnego podziału ze względu na sposób renderowania czy też możliwość interakcji gracza (której powyższa definicja nie przewiduje). W ramach materiału pre-renderowanego gracz obserwuje de facto odtwarzany plik wideo, który mógł być wyprodukowany w dowolny sposób. Typowy silnik gry nie jest w tym momencie używany, a filmik jest prezentowany w pewnego rodzaju odtwarzaczu multimedialnym. W przypadku materiału renderowanego w czasie rzeczywistym wykorzystywany jest silnik gry oraz modele/tekstury występujące podczas rozgrywki. Zapewnia to zdecydowanie płynniejsze przechodzenie pomiędzy momentami nieinteraktywnymi i interaktywnymi oraz prowadzi do większej spójności wizualnej. Wymaganie użycia sprzętu komputerowego może prowadzić do pewnego ograniczenia jakości czy też wykorzystywanych technik (jak np. symulacji fizycznych). Klasycznie cut scenki nie są interaktywne a gracz jest jedynie pasywnym odbiorcą. W niektórych tytułach możemy jednak natknąć się na metodę QTE (z ang. *quick time event*), w ramach której podczas odgrywania danej scenki wyświetlają się ikonki (potencjalnie wraz z instrukcjami) symbolizujące przyciski do wciśnięcia przez gracza. Wymusza to na nim uwagę przy oglądaniu materiałów a dodatkowo wprowadza mechanizm stresowy, ponieważ często błędnie wykonane sekwencje pociągają za sobą konsekwencje fabularne (np. śmierć danej postaci). Cut-scenki w wersji klasycznej jak i z uwzględnieniem QTE zostały zaprezentowane na rysunku 1.18.



(a) Normalna cut-scenka

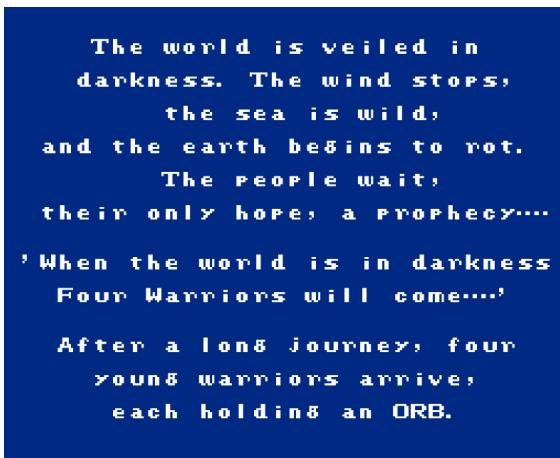


(b) Cut-scenka z QTE

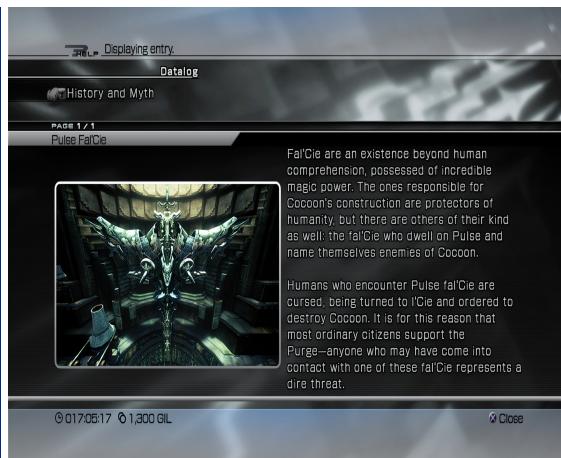
Rys. 1.18: The Dark Pictures Anthology: Man of Medan (2019) - Supermassive Games

Tekst

W przypadku tekstu możemy mówić o pełnoprawnych blokach tekstowych (Rys. 1.19), elementach interfejsu czy też o odpowiednich komunikatach pojawiających się na ekranie. Jest to oczywiście bardzo prosta forma przekazu fabularnego, wzorująca się na szeroko pojętej literaturze. Może być wykorzystany jako główny sposób prowadzenia opowieści lub jako element pomocniczy, często robiący fabularnie sens (np. w "Final Fantasy XIII" postacie naturalnie prowadzą między sobą dialog, ale informacje o świecie skryte w formie notatek znajdowanych przez gracza w trakcie rozgrywki).



(a) Final Fantasy I (1987)

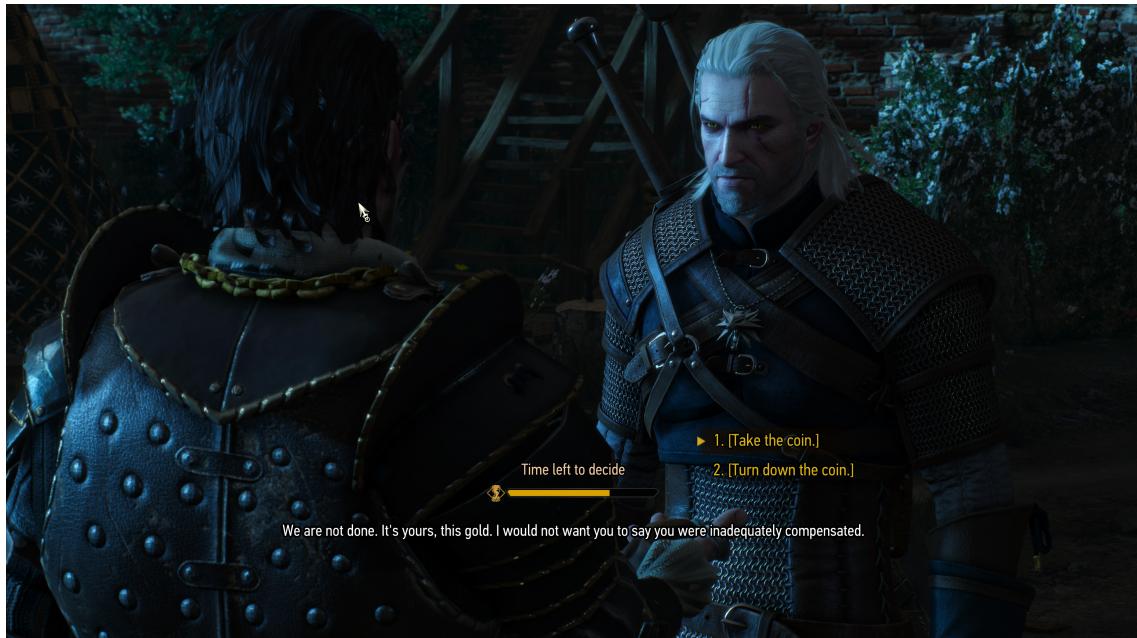


(b) Final Fantasy XIII (2009)

Rys. 1.19: Bloki tekstu służące do przedstawienia fabuły

Dialogi (głównie z NPC)

Dialog jako forma przedstawienia narracji może stanowić połączenie tekstu, dźwięku, animacji i interaktywności. Tekst może być obecny w formie napisów pomocniczych do kwestii wypowiadanych przez postacie, ale i również prezentuje możliwości do wyboru dostępne dla gracza. Współcześnie wiele gier nadaje głos swoim postaciom przy pomocy aktorów głosowych. Tak jak w przypadku cut scenek w czasie rzeczywistym, postacie najczęściej są animowane podczas dialogu, korzystając z silnika gry. Interaktywność podczas dialogu może wyłaniać się w formie przeklikiwania kolejnych kwestii (by dać graczowi czas na przeczytanie / zastanowienie się) lub poprzez możliwość wyboru kolejnych kwestii. W zależności od tytułu niektóre konwersacje mogą przypominać strukturę łańcucha pereł (gdzie rozmowa i tak kończy się w ten sam sposób) a niektóre formę rozgałęziającą się (gdzie odpowiedni wybór może nieść za sobą dalsze konsekwencje fabularne) [Patrz 1.2.1]. Więcej o systemach dialogowych wspomniane jest w sekcji 1.3.1.



Rys. 1.20: Dialog z NPC - "Wiedźmin 3" (2015)

Poprzez świat gry (audio-wizualne)

Najciekawszym a zarazem najrzadszym[30] sposobem prezentowania fabuły jest opowiadanie za pomocą samego świata gry. Mowa tu o obiektach i ich umiejscowieniu, tekstrach, ścieżkach dźwiękowych czy wszelkich innych widocznych lub słyszalnych elementach świata. Jedną z serii gier, która bazuje na tym koncepcie i cieszy się ogromną popularnością, jest "Dark Souls". Cut scenki są bardzo sporadyczne, przeważnie na początku i końcu rozgrywki oraz prezentujące starcie z trudnymi przeciwnikami ("bossami"). Spotykane postacie NPC są bardzo enigmatyczne, nie zadają graczowi wprost zadań do wykonania i

posiadają tylko kilka zapętlających się kwestii dialogowych. Tekstowy opis dotyczy głównie znajdowanych przedmiotów i zawiera szcątkowe informacje dotyczące ogólnopojętego świata gry. W ten sposób gracz musi samemu układać fabułę, na podstawie znajdowanych skrawków wiedzy. Dodatkowo, muzyka występuje tylko w ramach pojedynczych lokacji albo w przypadku starć z "*bossami*". W ten sposób autorzy podkreślają wagę tego co jest właśnie prezentowane na ekranie i starają się wywołać u gracza pewne emocje.

1.3. SYSTEMY DIALOGOWE W GRACH KOMPUTEROWYCH

Praca dotyczy wykorzystania sztucznej inteligencji do tworzenia angażującej narracji, a jest to realizowane właśnie stworzenie nowatorskiego systemu dialogowego opierającego się na dużych modelach językowych. Dlatego też warto prześledzić istniejące do tej pory systemy dialogowe spotykane w grach.

1.3.1. Popularne systemy dialogowe

Na podstawie dokonanego przeglądu tytułów można wyciągnąć pewne elementy wspólne i ubrać je w następujące kategorie: ze względu na formę, ze względu na precyzję i ze względu na wykorzystanie dodatkowych elementów. W ramach formy systemów dialogowych wyróżnione zostały: te, które ograniczają możliwości decyzyjne użytkownika; typ pudełkowy (box) oraz typ kołowy. Precyzja w tym przypadku oznacza dokładność pokrycia opcji przedstawionych graczowi z faktycznymi kwestiami wypowiadanymi przez postać. Jako dodatkowe elementy rozumiane są parametry takie jak czas, statystyki gracza czy ton wypowiedzi, które urozmaicają typowe dialogi.

Bez wyboru

W tym przypadku mowa o dialogach, które nie posiadają żadnej formy interaktywności (poza ewentualnym przewijaniem do kolejnych kwestii). Służą do przedstawienia narracji zaplanowanej przez producentów w imersywnej postaci — gracz bowiem jest swoistego rodzaju obserwatorem rozmowy. Oczywiście, nawet w interaktywnych systemach trudno sobie wyobrazić sytuację by gracz na każdą kwestię wypowiadaną przez NPC dokonywał wyboru odpowiedzi. W związku z tym można mówić, że ta forma dialogu jest obecna przynajmniej częściowo w każdym systemie.

Pudełkowy (box)

Klasyczną formą prezentowania dialogu jest podejście pudełkowe, gdzie tekst wyświetlny jest w specjalnym prostokącie widocznym na ekranie. Jest to sposób znany przede wszystkim ze starszych tytułów choć wcale nie zapomniany współcześnie. W ramach możliwych do podjęcie przez gracza decyzji wyświetlana jest lista opcji z odpowiednim wskaźnikiem, która opcja jest aktualnie podświetlona (Patrz rys. 1.21).



Rys. 1.21: Final Fantasy VII (1997)

Kołowy

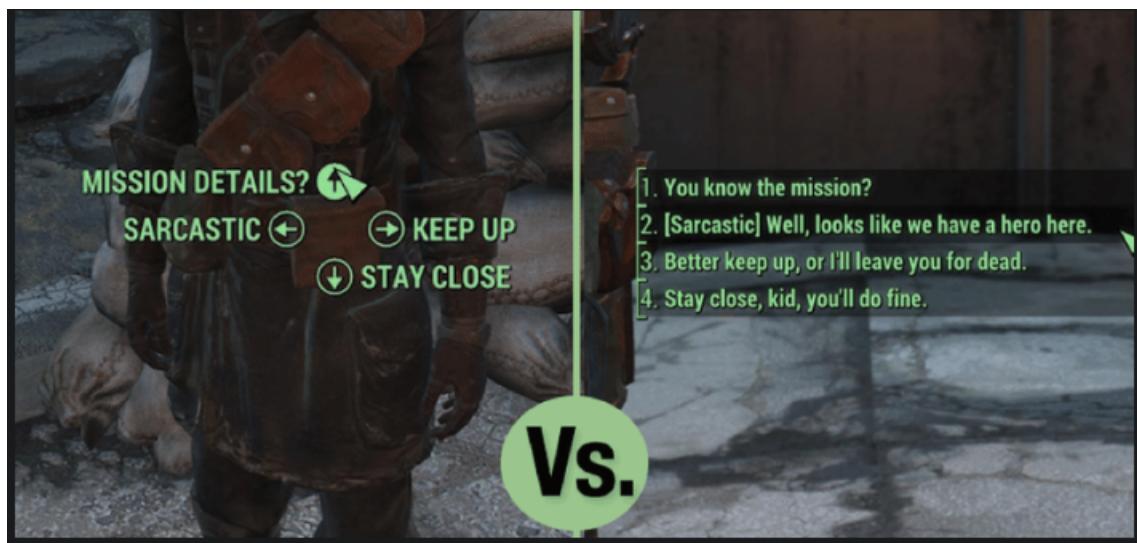
W ramach tego systemu mowa o klasycznym wyświetaniu napisów wspomagających dialog (tzn. mowa o transkrypcji kwestii wypowiadanych przez postacie, a dokładniej przez aktorów głosowych). Pojęcie koła pojawia się w momencie podejmowania przez gracza decyzji gdzie opcje ułożone w ogólnym rozumieniu w okręgu (Patrz rys. 1.22). Może to być podejście głównie znane z gier wspierających konsole, ze względu na analogowe gałki w kontrolerach, za pomocą których łatwo wybrać odpowiednią pozycję.



Rys. 1.22: Life is Strange (2015)

Precyjne / nieprecyjne

Jak wspomniano na początku sekcji, precyza określa pokrycie wyświetlanych opcji dialogowych z faktycznymi kwestiami wypowiadanymi przez postać. Niektóre tytuły są krytykowane właśnie za niezrozumiałe czy też nieintuicyjne wybory stawiane przed graczem. Przykładowo, w ramach gry "Fallout 4" mamy do czynienia z bardzo krótkimi 1-3 słownymi komunikatami, które nie oddają do końca tonu i intencji wypowiedzi. Społeczność fanowska utworzyła nawet modyfikację do gry, która zamienia kołowe i lakoniczne opcje na listę wyborów w formie zdań (Patrz rys. 1.23).



Rys. 1.23: Fallout 4 (2015) + wersja zmodyfikowana[10]

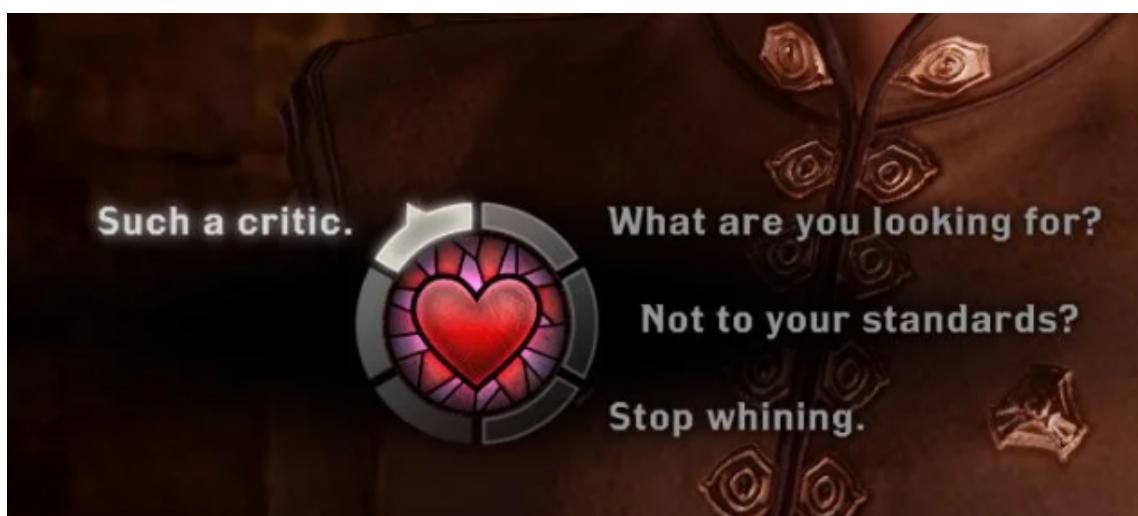
Wykorzystujące emocje

Systemy dialogowe mogą dodatkowo zawierać informacje o nacechowaniu emocjonalnym wypowiedzi. W grze "Dragon Age: Inquisition" można zaobserwować odpowiednie ikony, informujące grającego o tym w jaki sposób sterowana przez gracza postać wypowie daną kwestię. Wycinek ikon wraz z ich opisami został przedstawiony na rysunku 1.24.

	Stoic	Remain matter-of-fact about the subject at hand. Speak with self-control.
	Pleased	Express agreement with or happiness about the current conversation point.
	Sad	Express sadness over the current conversation point.
	Mad	Express your anger about what's being discussed.
	Confused	Indicate confusion about the current line of conversation.
	Surprised	Has the conversation astounded you in some way? Express your surprise.
	Anxious	Expresses tension, fears, anxiety. <i>Note: This isn't included in the manual, so the description is somewhat vague.</i>
	Attack	Enough conversation. It's time to get aggressive.

Rys. 1.24: Fragment spisu ikon dialogowych - "Dragon Age: Inquisition" (2014)[9]

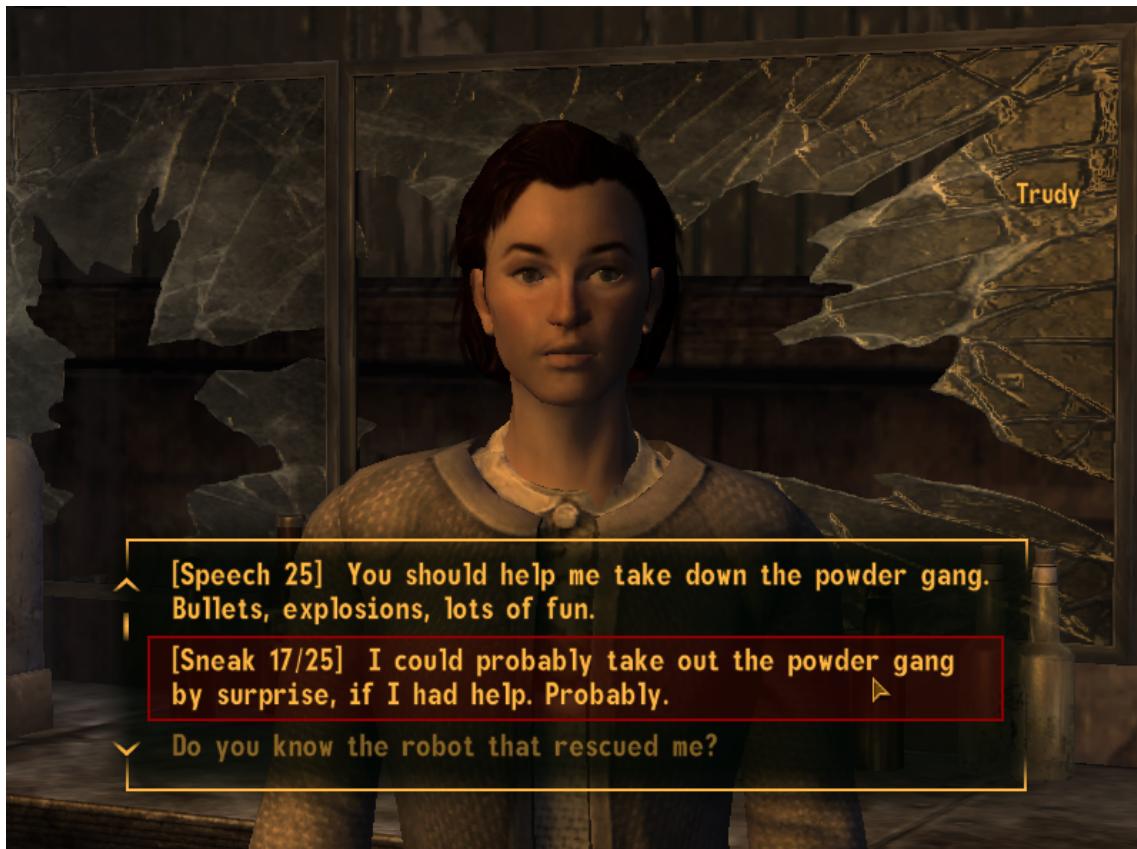
Ikony te są wyświetlane po najechaniu na odpowiednią opcję w momencie podejmowania decyzji (Patrz rys. 1.25).



Rys. 1.25: Przykład dialogu z wykorzystaniem ikony emocji - "Dragon Age: Inquisition" (2014)

Wykorzystujące statystyki

Niektóre tytuły, zwłaszcza te z gatunku RPG (ang. *role-playing game*) pozwalają rozwijać statystyki czy też atrybuty postaci (np. siła, charyzma). W takich grach możemy napotkać się na system dialogowy, w którym to pewne opcje są ograniczone czy też zablokowane ze względu na poziom statystyk postaci sterowanej przez gracza. Przykładowo, w "Fallout: New Vegas" decyzje dialogowe a co za tym idzie i fabularne, mogą ograniczać grającego do konkretnych rozwiązań (Patrz rys. 1.26).



Rys. 1.26: Fallout: New Vegas (2010)

Wykorzystujące czas

Spotykaną też czasami formą występującą w dialogach jest ograniczenie czasowe na podjęcie decyzji narzucone na graczącego przez grę. Jest to technika poniekąd inspirująca się metodą QTE znaną z cut scenek (Patrz sekcja 1.2.2). Rozwiążanie takie możemy znaleźć w grze "Wiedźmin 3" (Patrz rys. 1.27). Jeśli gracz nie podejmie decyzji w określonym czasie to albo kończy się to automatycznym wyborem jednej z dostępnych opcji albo występuje de facto "*trzecia opcja*".



Rys. 1.27: Wiedźmin 3 (2015)

1.3.2. Interaktywna fikcja - system poleceń

Innego rodzaju systemem dialogowym — a nawet i osobnym gatunkiem gier komputerowych — jest tak zwana *interaktywna fikcja*. Jest ona pewnego rodzaju oprogramowaniem symulującym środowisko, w którym to gracz używa wyłącznie komend tekstowych do poruszania się czy wpływania na to środowisko[14]. Według Nick'a Montfort'a pojęcie to może być utożsamione z "przygodami tekstowymi" czy prościej "grami tekstowymi"[23]. Z perspektywy nauczania maszynowego można uznać, że tego rodzaju gry zawierają w sobie elementy przetwarzania języka naturalnego (z ang. *NLP - natural language processing*) jak i sekwencyjnego podejmowania decyzji[11]. Przykład rozgrywki zostanie zaprezentowany w oparciu o polski tytuł "Otcpłan" (1999).

Po rozpoczęciu rozgrywki gracz znajduje się interaktywnym świecie gry, w którym różne kolory tekstu mają specyficzne znaczenia, wskazujące na różne elementy rozgrywki. Błękitny tekst podobny do <21hp 104m 112mv 70exp> oznacza status gracza, a sama gra oczekuje na polecenie. Zielony tekst oznacza lokacje, w których znajduje się gracz, a różowy

tekst wskazuje możliwe wyjścia z danego miejsca. Krótkie opisy dają graczowi dodatkowe informacje o otoczeniu. Całość tworzy spójną, interaktywną fabułę, w której gracz może podejmować decyzje i eksplorować świat gry. Powyższe opisy są do zaobserwowania na rysunku 1.28.

```

Lothar zamierza juz odejsc, ale jednak odwraca sie jeszcze na chwile - "Dobra, zrobiny tak. Jesli nadal twierdzisz, ze nie wiesz kim jestes to byc moze bede mogl ci pomoc. Ale nie tu i nie teraz i zapewne nie za darmo. Obejrzyj sobie miasto, zdobadz troche doswiadczenia w walce - i tak bedzie ci potrzebne. Bede wiedzial o twoich postepach i skontakuje sie z tobą wkrótce. Te mikstury tez ci sie moga przydac. Powodzenia!"  

Mag Lothar odchodzi stad na zachod.  

<21hp 104m 112mv 78exp>  

Skrzyzowanie  

Wyjścia: east west north south  

Stoisz na skrzyżowaniu ulic, bedacym jednocześnie dużym placem. Na zachód przechodzi on w niewielki targ, za którym stoi świątynia. Na południe niewielka uliczka biegnie wzdłuż muru. Na północ plac kończy się przy rzece, za którą stoi jakas rezydencja.  

<21hp 104m 111mv 78exp>  

Plac targowy  

Wyjścia: east west north south  

Niesamowity tłok! Ludzie pchają się na Ciebie i próbują sprzedac co mają; rzedy straganów ciągną się z obu stron placu. Na zachód widać kolosalny budynek: świątynię jakiegoś bohatera.  

<21hp 104m 111mv 78exp>

```

Rys. 1.28: Ogólny wygląd rozgrywki w "Ottchłani"

Dialogi występujące w "Ottchłani" z postaciami NPC prowadzone są w formie listy wybieranych kwestii z perfekcyjną precyją (Patrz sekcja 1.3.1), tzn. kwestia wypowiadana jest w takiej formie, w jakiej występuje ona w menu. Sama konwersacja rozpoczynana jest oczywiście za pomocą odpowiedniej komendy, a zakończona może być przez odpowiedni wybór użytkownika lub po odpowiedzi od postaci NPC (co widać na rysunku 1.29).

```

Dziedzinec  

Wyjścia: north south  

Stoisz na szerokim dziedzincu. To tutaj Caranel organizuje, wyposaża i stąd wysyła swoje wsparcie karawany do nienanich krajów w celu sprowadzenia tajemniczych i fantastycznych towarów. Na wschód widzisz stajnie, a na zachód kuznie. Z kolei na południe widzisz wyjście do miasta, a na północ wejście do rezydencji.  

Strażnik mostu strzeże wejścia do rezydencji Caranela.  

<21hp 104m 97mv 66exp>zagadaj strażnik  

Strażnik mostu poprawia pancerz.  

"Witam."  

1> "Czyja to rezydencja?"  

2> "Czego Pan tu pilnuje?"  

3> "Witam... i zegnam."  

1  

"Czyja to rezydencja?"  

"Rezydencja ta należy do kupca - Hansa Caranela. Jest to najbogatszy kupiec w Mantarze. Toteż uważaj bym Cię nie przylapał na próbce zwiedzenia czegos."  

1> "Gdzie go znajdę?"  

1  

"Gdzie go znajdę?"  

"Powinieneś być na piętrze w swoim gabinecie. Ewentualnie na dole w bibliotece."  

<21hp 104m 97mv 66exp>

```

Rys. 1.29: Dialog w "Ottchłani"

Gra oferuje pewnego rodzaju katalog dostępnych dla gracza poleceń, które pogrupowane są w odpowiednie kategorie tematyczne. Aby uzyskać listę wystarczy wydać polecenie

"pomoc". Mimo braku szczegółowych opisów, komendy zostały zaprojektowane w formie dość intuicyjnych poleceń jako słowa z języka polskiego. Pełna lista dostępnych komend jest widoczna na rysunku 1.30.

```
<21hp 104m 112mv 66exp>pomoc
PORUSZANIE SIE:
    east, west, north, south, up, down
TRANSAKCJE:
    kup, sprzedaj, lista, zamow
PRZEDMIOTY:
    wez, zostaw, zaloz, zdejmij,
    zazyj, zjedz, wypij, poswiec
PRZECIWNICY:
    zabij, kontra, spojrza
INFORMACJE O GRACZU
    dane, umiejetnosci, dzialanie
    ekwipunek, uzbrojenie, czas
POLECENIA GRY
    zapisz, wyjdź, alias, unalias,
    /konfiguracja </przyporzadkuj ! >
CZYNNOŚCI:
    spij, odpoczywaj, spojrza,
    rozejrzyj, czytaj, cwicz, trenuj,
    otworz, zamknij, pij, czaruj
TEMATY OGÓLNE
    nowy gracz, walka, swiat, przedmioty, chodzenie, moby, strzelanie
    choroby, przygody, przeklenstwa, reputacja, kranki, dziennik
```

Rys. 1.30: Polecenia dostępne w "Ochłani"

W przypadku gdy grający ma problem ze zrozumieniem pewnych komend albo samego sposobu prowadzenia rozgrywki to może sięgnąć po samouczek dostępny na oficjalnym forum "Ochłani". Na rysunku 1.31 zauważać można szczegółowy opis komend związanych z kategorią eksplorowania świata gry.

Samouczek

■ autor: Weq » ndz wrz 20, 2015 22:44

0) Interfejs
Gra Ochłani jest grą tekstową, co oznacza, że wszystkie komendy wydawane są w wierszu poleceni i wyrażają się po prostu przez odpowiednie komendy (np. "zjedz", "wypij", "otwórz"). Prawie wszystkie komendy są dostępne pod hasłem "**pomoc**", jeżeli komenda nie jest rozpoznana przez grę, otrzymujemy znaczące "Hmmm?".
Komendy można wydawać, skracając je, oraz bez polskich liter (gra je rozpozna).

1) Eksploracja
Najważniejsze komendy: east, west, north, south, otwórz, wejdź
Na początek parę słów odnośnie Świata. Podstawą "częścią" świata jest lokacja. Każda lokacja ma nazwę, własny opis, a także oznaczenie wyjść. Dodatkowo może się znajdować w niej mob (npc), oraz przedmioty. Niektóre rzeczy (np pojemniki, dzwignie) nie są wypisane wprost, ale są częścią opisu. Przeczytaj też "pomoc lokacji".
Poruszanie się po świecie Ochłani jest bardzo proste, wystarczy wpisać nazwę kierunku (po angielsku (1)), w którym chcemy się poruszyć, czyli : "north" (północ), "south" (południe), "east" (wschód) i "west" (zachód). Dodatkowo w niektórych lokacjach można poruszać się jeszcze w górę (up) i w dół (down). Zamiast pełnych nazw najwygodniej używać pierwszych liter. Nie zawsze możliwe wyjście jest wymienione (wyjścia ukryte).
Dodatkowo, jeśli w lokacji znajduje się portal (lub coś w tym stylu), trzeba mieć w pamięci komendę "**wejdź**". Jeśli chcemy jeszcze raz przeczytać opis lokacji, wystarczy wpisać "**spójrz**".
Gra nie posiada mapy, ale można takową znaleźć jako przedmiot, choć pewnie nie będzie zbyt dokładna.

Rys. 1.31: Fragment samouczka "Ochłani"

W pracach nad "Ochłanią" zaangażowane były dwie osoby: Grzegorz 'Weq' Nowak oraz Krzysztof 'Hoborg' Ciesielski. Zaczęli oni pracę nad grą w 1998 roku mając odpowiednio 16 i 14 lat[21]. Po kilku latach Krzysztof Ciesielski odszedł od projektu. Ostatecznie, Grzegorzowi Nowakowi udało się po kilkunastu latach doprowadzić tytuł do wersji finalnej dostępnej za darmo na stronie <https://www.ochlan.pl>.

2. SPOSÓBY GENEROWANIA NARRACJI

W poprzednich sekcjach omówiona została historia oraz wykorzystywane w grach komputerowych rodzaje narracji. W każdym z tych elementów istnieje jeden element wspólny: to ludzie odpowiadają za tworzenie narracji, odpowiednie jej planowanie i prezentowanie odbiorcy. Związane z tym są oczywiście olbrzymie koszty oraz duży nakład czasu poświęcony przez pracowników. Biorąc pod uwagę terminy stale goniące producentów gier, nic dziwnego, że pewne zaplanowane fragmenty fabularne nie znajdują miejsca w końcowym produkcie. Z tego względu, patrząc na postępujący rozwój technologiczny, nasuwa się pytanie — czy da się zarządzać narracją w grze komputerowej w sposób automatyczny? W ramach tej sekcji przedstawione zostaną znane rozwiązania z dziedziny sztucznej inteligencji pomagające w kreowaniu fabuły oraz nakreślony zostanie potencjał w tej sferze dużych modeli językowych.

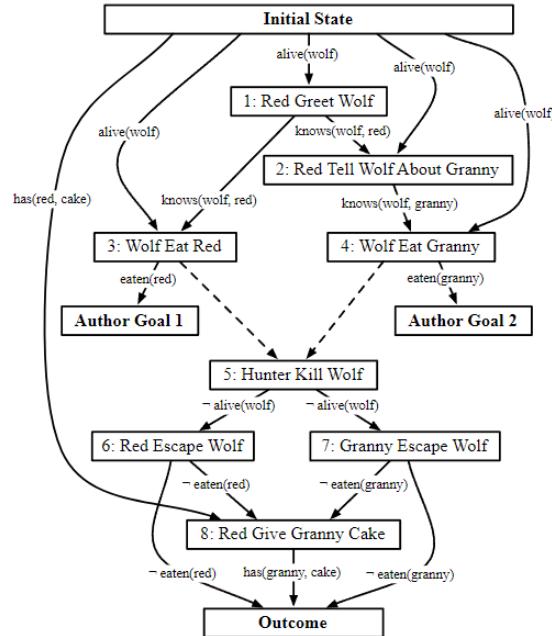
2.1. WYKORZYSTANIE ALGORYTMÓW SZTUCZNEJ INTELIGENCJI DO KREOWANIA NARRACJI

Mówiąc o ogólnym wykorzystaniu algorytmów sztucznej inteligencji można cofnąć się do bardzo prostych rozwiązań wykorzystanych np. w "Pong" (Patrz sekcja 1.1.2), do technik generowania proceduralnego (zwłaszcza poziomów) czy też do systemów rankingowych (np. system TrueSkill). Jako, że nie są to metody stricte związane z narracją to nie zostaną one opisane bardziej szczegółowo. Przedstawione natomiast będą kluczowe obszary wykorzystywane w grach: częściowo-uporządkowane planowanie (ang. *POP* - partially-ordered planning), modelowanie doświadczeń gracza (ang. *PEM* - player experience modelling), przetwarzanie języka naturalnego (ang. *NLP* - natural language processing), postać niegrywalna (ang. *NPC* - non-playable character), proces decyzyjny Markowa (ang. *MDP* - Markov decision process).

POP (Partially-ordered planning)

Planowanie częściowo-uporządkowane jest skierowanym grafem acyklicznym, gdzie węzły są operacjami (inaczej nazywanymi akcjami), które po wywołaniu zmieniają stan świata. Krawędzie przedstawiają relacje przyczynowe i czasowe pomiędzy akcjami. Powiązanie przyczynowe $a_i \rightarrow^c a_j$ oznacza, że wykonanie akcji a_i zmieni stan warunku c na prawdziwy w świecie fabuły, a co za tym idzie akcja a_j zależna od tego warunku będzie możliwa do wykonania. Powiązanie czasowe przedstawia ograniczenie porządkowe

pomiędzy operacjami, gdzie jedna operacja musi być wykonana przed inną[29]. Przykładowa struktura fabularna zrealizowana za pomocą planowania częściowo-uporządkowanego została przedstawiona na rysunku 2.1.



Rys. 2.1: Fabuła "Czerwonego Kapturka" zapisana za pomocą POP

Za pomocą tej techniki kreować można rozbudowane plany fabularne, które mogą ulegać zmianie na podstawie akcji podejmowanych przez gracza czy zmian zachodzących w świecie gry. Odpowiednie algorytmy przeszukiwania nazywane "*plannerami*" rozwiążają problem planowania, tzn. mając dany stan świata, pewne atomowe akcje możliwe do wykonania przez gracza oraz założony cel, znajdują odpowiednią sekwencję operacji, które doprowadzą do osiągnięcia tegoż celu[29].

Podstawowym problemem pojawiającym się przy wykorzystaniu metody POP jest to, że zarówno gracz jak i potencjalnie inne niegrywalne postacie, mogą być zdolne do wywołania akcji w świecie gry, która zagraża dalszemu przebiegowi fabularnemu zgodnego z planem[19]. Wtedy stosowane są odpowiednie techniki naprawcze, które prowadzą do mniej lub bardziej doskonałych rozwiązań.

PEM (Player experience modelling)

Modelowanie doświadczeń gracza polega na zbieraniu danych behawioralnych czy też wydajnościowych (punkty, czas, decyzje) ze względu na rozgrywkę za pomocą wielu modalności: mowy gracza, obrazów (śledzenie ruchów ciała, mimiki twarzy, gałek ocznych) czy też sygnałów fizjologicznych (puls czy przewodność skóry). Pomiar sygnałów fizjologicznych jest oczywiście problematyczny ze względu na wykorzystanie dodatkowego sprzętu a zarazem inwazyjność przeszkadzającą w swobodnej rozgrywce [32].

W ramach tego obszaru sztuczna inteligencja objawia się zazwyczaj pod postacią sieci neuronowych czy też drzew decyzyjnych, które pozwalają dokonywać klasyfikacji w zakresie[32]:

- rozpoznawania emocji grającego - w ramach anagażujących systemów dialogowych
- balansowania rozgrywki - tak by gracz nie odczuwał frustracji ze względu na zbyt wysoki poziom trudności a zarazem by nie doznał nudy ze względu na zbyt prostą rozgrywkę
- oceniania umiejętności gracza - do prowadzenia badań w sposób ukryty

NLP (Natural language processing)

Przetwarzanie języka naturalnego jest dziedziną w obrębie sztucznej inteligencji, która zajmuje się zrozumieniem, interpretacją i manipulacją ludzkiego języka[32]. W ramach gier komputerowych pozwala to graczowi poruszać się po świecie czy też komunikować z innymi postaciami NPC w sposób zarówno naturalny (zamiast wchodzenia w interakcję z odpowiednimi interfejsami) jak i dość otwarty (na tyle na ile dany system jest w stanie przetwarzać odpowiednie frazy).

Z podstaw przetwarzania języka naturalnego korzystały tytuły realizowane w konwencji interaktywnej fikcji, takie jak "Otochłań" (Patrz sekcja 1.3.2). Pozwalają one graczowi za pomocą określonego zestawu komend poruszać się i wchodzić w interakcję z całym światem gry.

Związane z tą dziedziną są duże modele językowe, które potrafią operować językiem naturalnym. Ich bardziej szczegółowy opis będzie przedstawiony w sekcji 2.2.

NPC (Non-playable character)

Postacie niegrywalne to wszelkie jednostki czy postacie, z którymi gracz może wchodzić w interakcję lub dostrzegać ich autonomiczne poruszanie w świecie. Mogą to być statyczne byty, które usytuowane są w jednym określonym miejscu i mają na celu dawać grajecemu zadania czy też informacje o świecie gry. Z drugiej strony wszelkie jednostki, z którymi gracz walczy również podlegają pod tę definicję. Forma postaci niegrywalnych, tak jak i w literaturze, nie jest istotna, gdyż bohaterem może być zarówno człowiek jak i zantropomorfizowane zwierzę czy rzecz.

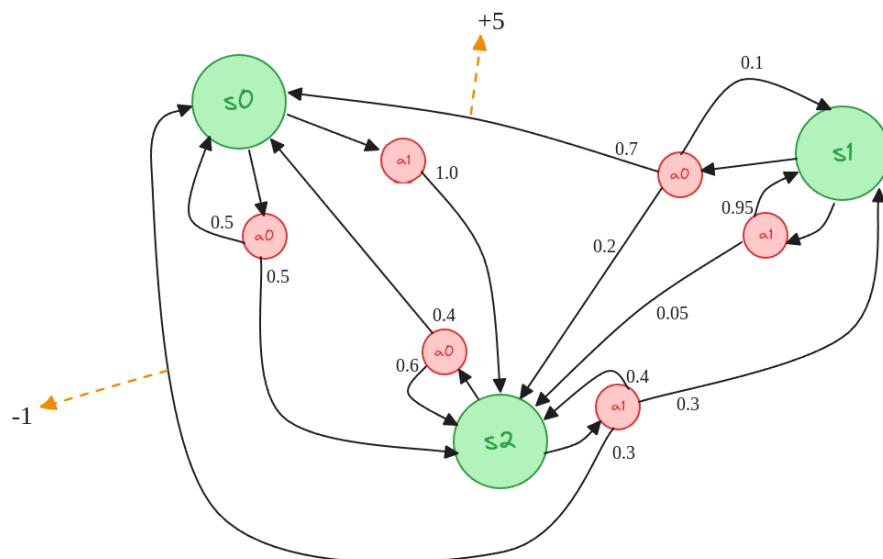
NPC pojawiły się w grach z początkiem lat 90-tych. Oparte były przede wszystkim na zpredefiniowanych skryptach i drzewach decyzyjnych[26] (współcześnie nadal wiele postaci jest opartych o te rozwiązania) [6].

Wraz z rozwojem technologicznym twórcy gry mają możliwość przeznaczenia więcej mocy obliczeniowej na wiarygodne dla gracza postacie NPC. Związane jest to z dokładniejszymi modelami i animacjami postaci ale również z bardziej zaawansowanymi wzorcami zachowania. Producenci zaczęli wykorzystywać od lat 2010 w tym celu techniki uczenia

maszynowego oraz głębokiego nauczania [6]. Pozwala to przeciwnikom (czy też sprzymierzeńcom) gracza dostosowywać się do sposobu prowadzenia przez niego rozgrywki. Współczesne tytuły takie jak "Read Dead Redemption 2" czy "The Last of Us Part II" wykorzystują głębokie sieci neuronowe do budowania postaci NPC[6].

MDP (Markov decision process)

Łańcuch Markowa jest stochastycznym modelem opisującym ciąg możliwych zdarzeń, w którym prawdopodobieństwo każdego zdarzenia zależy jedynie od wyniku poprzedniego. Rozróżniane one są dodatkowo ze względu na dyskretne lub ciągłe momenty czasowe, w których następuje zmiana stanów. Proces decyzyjny Markowa (MDP) jest zasadniczo rozszerzeniem łańcuchów Markowa — różnicę stanowi dodanie akcji (które pozwalają na podejmowanie decyzji) oraz nagród otrzymywanych po przejściu z jednego stanu w drugi. Jeżeli dla każdego stanu istnieje tylko jedna akcja i wszystkie nagrody są jednakowe, to proces decyzyjny Markowa upraszcza się do łańcucha Markowa.



Rys. 2.2: Przykładowy proces decyzyjny Markowa

Na powyższym rysunku przedstawiony został przykładowy proces decyzyjny Markowa z trzema stanami (zielone kółka), dwiema akcjami (czerwone kółka) oraz dwiema nagrodami (pomarańczowe strzałki). Model MDP może być wykorzystany w ramach generowania treści gier komputerowych do budowania zadań dla gracza, przykładowo dla gry kucharskiej baza składników może być ułożona w odpowiednie łańcuchy Markowa, tak by bardziej pasujące do siebie składniki miały większe prawdopodobieństwo bycia wspólnie wybranym [5].

Proces może być dodatkowo wykorzystany w proceduralnym generowaniu treści czy świata dla gracza. Zakładając, że to gracz jest "*agentem*" podejmującym akcje w świecie

gry (które są jawnie zdefiniowane przez twórców), doprowadza on do zmiany stanu świata co jest idealnie reprezentowane przez proces decyzyjny Markowa[28].

Przykład wykorzystania - "Façade"

"Façade" (2005) to tytuł, który próbował złamać znany do tamtej pory podział na narrację liniową czy też rozgałęzającą się, na rzecz w pełni interaktywnej historii kontrolowanej przez gracza. Jest to trójwymiarowa gra czasu rzeczywistego przedstawiona w formie jednego aktu. Narracja koncentruje się wokół Grace i Tripa, małżeństwa po trzydziestce, które zaprasza gracza na drinka. Gracz nie wie, że ich małżeństwo znajduje się na grząskim gruncie, a tego wieczoru wszystkie ich małżeńskie problemy wyjdą na jaw. To, w jaki sposób ich związek się rozpadnie, a także ostateczny związek gracza z Grace i Tripem, zależy od interakcji gracza ze światem. Gracz angażuje się poprzez poruszanie się po środowisku, manipulowanie przedmiotami i, co najważniejsze, poprzez dialogi w języku naturalnym[22].

Twórcy na potrzebę "Façade" stworzyli własny język ABL ("A Behavior Language" - z ang. język zachowań). Stanowi on pewnego rodzaju połączenie drzew decyzyjnych i procesów decyzyjnych Markowa wspomnianych wcześniej. Programy ABL są zorganizowane jako zbiór zachowań, które mogą być procesowane sekwencyjnie lub równolegle. Dodatkowo, zachowania mogą mieć narzucone określone wymagania by mogły one się odbyć, a w zależności od powodzenia lub porażki w realizacji zachowania świat gry może zostać odpowiednio zmieniony.

2.2. WYKORZYSTANIE DUŻYCH MODELI JĘZYKOWYCH (LLM) DO KREOWANIA NARRACJI

W ostatnich latach coraz większa liczba prac naukowych oraz rozwiązań komercyjnych skupia się wokół tzw. dużych modeli językowych. Okazuje się bowiem, że wprowadzając optymalizacje w sferze oprogramowania czy też po prostu przeznaczając więcej mocy obliczeniowej na wyuczenie takowych modeli, zyskuje się adekwatnie większą "skuteczność" jeżeli chodzi o wykonywane przez nie zadania z zakresu przetwarzania języka naturalnego (i nie tylko!).

W związku z tym nasuwa się następujące pytanie — czy można wykorzystać **aktualne** duże modele językowe do skutecznego kreowania narracji? W ramach tej sekcji przedstawiona zostanie dokładniejsza charakterystyka tych modeli a zarazem istniejące próby ich wykorzystania.

Definicja i charakterystyka dużych modeli językowych

Duże modele językowe są modelami nauczania maszynowego, które są w stanie wykonywać zadania z dziedziny przetwarzania języka naturalnego, takie jak generowanie tekstu,

tłumaczenie tekstu z jednego języka na inny czy prowadzenie rozmowy z człowiekiem w sposób konwersacyjny[34]. Pojęcie "duże" oznacza w tym przypadku zarówno skalę samych modeli (miliardy czy biliony parametrów) jak i wymaganą do ich wyuczenia ilość danych. Pojęcie LLM skoncentrowane jest zasadniczo wokół funkcjonalności danego modelu, niezależnie od przyjętej do jego wyuczenia architektury.

Z uwagi na fakt, że zdecydowana większość modeli oparta jest na architekturze transformera to zostanie ona krótko nakreślona. Dane podawane na wejście do modelu mogą być różnych modalności np. tekst, audio czy obraz. Następnie, dane te dzielone są na tzw. żetony (ang. *tokens*) - fragmenty tekstu, obrazu lub dźwięku. Każdy żeton jest kodowany jako wielowymiarowy wektor liczbowy, odzwierciedlający jego semantyczne właściwości. Następnie sekwencja wektorów przechodzi przez blok uwagi (ang. *attention*), gdzie wektory wzajemnie się aktualizują (ponieważ np. znaczenie słowa zależy od kontekstu, czyli od słów występujących dookoła niego). Ten proces pozwala na lepsze zrozumienie zależności pomiędzy elementami wejściowymi. Po bloku uwagi, wektory są przetwarzane przez perceptron wielowarstwowy (ang. *MLP - multi-layer perceptron*). Cykl bloków uwagi i MLP jest wielokrotnie powtarzany. Na końcu procesu otrzymujemy pojedynczą macierz, na podstawie której, poprzez operację softmax, generowany jest rozkład prawdopodobieństwa dla możliwych żetonów następujących po danym fragmencie wejściowym. Autorzy raportu technicznego dotyczącego GPT-4 zauważają, że:

"...Takie modele są ważnym obszarem badań, ponieważ mają potencjał do wykorzystania w szerokim zakresie zastosowań, takich jak systemy dialogowe"[25]

Można również wnioskować, że następne iteracje modeli będą coraz lepsze:

"Jednym z głównych celów rozwoju takich modeli jest poprawa ich zdolności do rozumienia i generowania tekstu w języku naturalnym, szczególnie w bardziej złożonych i znuansowanych scenariuszach. "[25]

Wymienione są jednak przez autorów ograniczenia aktualnych modeli językowych, które uniemożliwiają ich masową adopcję w wielu przestrzeniach:

"Pomimo swoich możliwości, GPT-4 ma podobne ograniczenia do wcześniejszych modeli GPT: nie jest w pełni niezawodny (np. może cierpieć na „halucynacje”), ma ograniczone okno kontekstowe i nie uczy się na podstawie doświadczenia. Należy zachować ostrożność podczas korzystania z wyników GPT-4, szczególnie w kontekstach, w których ważna jest niezawodność. "[25]

Pod pojęciem "halucynacji" rozumiane jest wymyślanie faktów, ponieważ modele te same w sobie nie posiadają żadnej podpiętej bazy wiedzy a generowany przez nie tekst wynika z modelu probabilistycznego (co nie znaczy, że w ogóle nie dysponują wiedzą). Wielkość okna kontekstowego oznacza na jakim rozmiarze danych / żetonów model jest w

stanie jednocześnie pracować (tzn. w przypadku bardzo długich konwersacji czy plików model może nie brać pod uwagę wczesnych informacji przy udzielaniu odpowiedzi).

Przykłady zastosowania dużych modeli językowych

W badaniu Schaap i innych (2023)[4] zbadano potencjał modeli BERT i GPT-2 do proceduralnego generowania zadań (questów) w grze QuestVille. Podejście to polegało na połączeniu dwóch modeli językowych: BERT i GPT-2 w sekwencji. Najpierw zdefiniowano krótkie zdania będące podstawowymi zadaniami, z jednym słowem zamaskowanym. Zdania te przesyłano do modelu BERT, który przewidywał najbardziej prawdopodobne słowa pasujące do kontekstu. Następnie losowo wybierano jedno ze zwróconych słów i umieszczało je w zdaniu, tworząc wejście dla modelu GPT-2. Model ten generował dodatkowy tekst narracyjny uzupełniający podstawowe zadanie o motywacje, opis sytuacji i wprowadzał elementy fabularne. Wyniki sugerują, że takie połączenie modeli ma potencjał do generowania angażujących zadań z narracją, bardziej wciągających niż podstawowe polecenia. Poprzedzanie wejść relacjami między postaciami niegrywalnymi w grze kierowało generowaną treść w odpowiedni kontekst. Jednak generowane zdania nie zawsze były w pełni spójne i odpowiednie. Autorzy uznali, że dalsze postępy w dziedzinie NLP, w tym pojawienie się nowszych, lepszych modeli, mogą doprowadzić do jeszcze lepszych rezultatów w przyszłości.[4]

Jak wskazują w swojej pracy Umbraško i Drury (2023)[31], duże modele językowe mogą odgrywać kluczową rolę w dynamicznym tworzeniu treści gry na podstawie wejściowych danych tekstowych i kontekstu rozgrywki. Opisany prototyp gry wykorzystuje interfejs API ChatGPT do generowania początkowej sceny narracyjnej oraz opisu przeciwników (orków) na podstawie wysłanego do modelu żądania w formacie JSON. Kluczowym elementem jest też dynamiczne tworzenie dalszych fragmentów narracji na podstawie działań gracza i stanu gry, w tym cech przypisanych wrogom (np. "płonący", "pijany"). Wprowadzanie tych cech jako danych wejściowych pozwala modelowi językowemu na bardziej kontekstowe dopasowanie narracji. Autorzy prototypu zdecydowali się na ograniczenie zakresu możliwych wyjść modelu (np. rodzaje broni wrogów) dla zachowania spójności z mechanicznymi elementami gry. Wskazuje to na konieczność znalezienia właściwej równowagi pomiędzy swobodą generatywną modelu a wymogami spójnego i czytelnego doświadczenia. Wyzwaniem opisanym w pracy było zmapowanie całego stanu gry, takiego jak pozycje postaci, jako danych wejściowych do modelu ChatGPT. Ostatecznie autorzy zdecydowali się na prostszy model, w którym tylko kluczowe elementy stanu gry są przekazywane do generowania narracji. Zaprezentowany prototyp stanowi ciekawą próbę praktycznego wykorzystania dużych modeli językowych do tworzenia angażującej, interaktywnej narracji w grze wideo w oparciu o działania gracza. Pokazuje zarówno obiecujące możliwości, jak i obecne ograniczenia takiego podejścia.[31]

W pracy "LARP: Language-Agent Role Play for Open-World Games" (2023)[34] auto-

rzy proponują wykorzystanie dużych modeli językowych jako podstawę pod generatywnych agentów występujących w środowiskach otwartych światów gier fabularnych. Kluczowym elementem architektury zaproponowanej w pracy jest kognitywna architektura agenta inspirowana psychologią poznawczą. Składa się ona z modułów odpowiedzialnych za pamięć długotrwałą, roboczą, przetwarzanie pamięci oraz podejmowanie decyzji. Pozwala to na symulację ludzkich procesów poznawczych, takich jak kodowanie, przechowywanie i przypominanie informacji z pamięci, a także rekonstrukcję zdarzeń oraz zapominanie. Integracja tych mechanizmów z LLM umożliwia agentowi prowadzenie spójnej, długoterminowej narracji w otwartym świecie gry. Kolejnym istotnym aspektem jest moduł interakcji ze środowiskiem, który przekłada decyzje agenta na konkretne działania w grze. Wykorzystuje on bibliotekę akcji podstawowych oraz generuje nowe sekwencje akcji przy użyciu dopasowanego LLM w celu realizacji zamierzonych celów agenta. Ten proces stale wzbogaca bibliotekę akcji o nowe schematy postępowania. Ponadto, praca zakłada implementację mechanizmu dopasowywania różnorodnych osobowości i perspektyw dla agentów za pomocą zbioru drobniejszych, wyspecjalizowanych modeli LLM. Pozwala to na generowanie narracji dostosowanych do zróżnicowanych tożsamości, stylów językowych i postaw bohaterów niegrywalnych. Podejście LARP łączy zaawansowane techniki przetwarzania języka naturalnego z inspiracjami z psychologii poznawczej w celu umożliwienia generowania spójnych, długoterminowych narracji agentów w bogatych, otwartych światach gier fabularnych[34].

W ramach pracy "Generative Agents: Interactive Simulacra of Human Behavior" (2023)[27] autorzy przedstawiają koncepcję generatywnych agentów, które wykorzystują duże modele językowe do symulacji wiarygodnych zachowań ludzkich. W porównaniu z poprzednią, ta prezentuje wnioski dotyczące faktycznej implementacji tej struktury. Architektura agenta opiera się na trzech głównych komponentach: strumieniu pamięci, refleksji i planowaniu. Strumień pamięci to moduł długoterminowej pamięci, który rejestruje doświadczenia agenta w języku naturalnym. Refleksja pozwala agentowi na wyciąganie wniosków na wyższym poziomie abstrakcji, co przekłada się na lepsze kierowanie jego zachowaniem. Planowanie to proces przekształcania tych wniosków i aktualnego środowiska w wysokopoziomowe plany działań, które są następnie realizowane przez agenta. Autorzy przeprowadzają dwie ewaluacje agentów generatywnych: kontrolowaną ewaluację indywidualnych zachowań oraz kompleksową ewaluację interakcji między agentami w otwartym środowisku przez dwa dni czasu gry. W ewaluacji technicznej wykorzystują metodę "wywiadu" z agentem w języku naturalnym, aby zbadać jego zdolność do pozostania w charakterze, pamiętania, planowania, reagowania i dokładnego refleksji. Porównując różne wersje systemu z ograniczonym dostępem do pamięci, refleksji i planowania, obserwując, że każdy z tych komponentów jest kluczowy dla wiarygodności zachowań agenta[27].

Jak widać obecnie trwają intensywne prace badawcze nad wykorzystaniem dużych modeli językowych, w dziedzinie tworzenia treści literackich oraz elementów gier. Jednym z

obiecujących zastosowań jest stworzenie w pełni autonomicznych, generatywnych agentów literackich, zdolnych do samodzielnego tworzenia spójnych i angażujących narracji. Rozwiązania te stopniowo znajdują swoje odzwierciedlenie również na rynku komercyjnym, czego przykładem są platformy takie jak Inworld.ai czy Nvidia Avatar Cloud Engine. Umożliwiają one twórcom gier, a nawet indywidualnym użytkownikom, korzystanie z zaawansowanych modeli językowych w celu generowania dialogów postaci, opisów środowisk czy całych wątków fabularnych. Chociaż wciąż istnieją liczne wyzwania natury technicznej i etycznej, rozwój tej technologii może znacząco zmienić oblicze branży rozrywkowej oraz procesów twórczych w dziedzinie literatury.

3. ZAANGAŻOWANIE GRACZA

By móc określić wpływ rozwiązań sztucznej inteligencji na angażującą narrację wewnątrz gry należy zdefiniować w jaki sposób określić i mierzyć można zaangażowanie gracza. W tej sekcji podjęta zostanie próba skategoryzowania zaangażowania i wybrania najodpowiedniejszej metody do wykorzystania w ramach eksperymentu.

3.1. DEFINICJE I RODZAJE ZAANGAŻOWANIA

Pomiar i ocena zaangażowania graczy w gry wideo jest niezwykle skomplikowanym zadaniem. Podstawową przeszkodą jest to, że zaangażowanie nie jest zjawiskiem jednowymiarowym[3], łatwym do uchwycenia[13]. Stanowi raczej złożony, subiektywny stan umysłu, na który składać się może wiele elementów. W literaturze naukowej poświęconej temu zagadnieniu można najczęściej spotkać się z terminami takimi jak: imersja, obecność, przepływ / trans, absorpcja psychologiczna czy dysocjacja. Podstawowym problemem jest fakt, że doświadczenia przeżywane podczas rozgrywki różnią się ze względu na gatunek gry, na stan emocjonalny / psychiczny grającego czy też ze względu na jego charakter[13].

Kolejnym kluczowym utrudnieniem jest fakt, iż zaangażowanie oparte jest na nieświadomych procesach poznawczych i emocjonalnych, do których trudno uzyskać introspekcjyjny dostęp[13]. Wymaganie od gracza, by w trakcie rozgrywki analizował i werbalizował swoje przeżycia, nieuchronnie zakłóca i niszczy stan umysłu będący celem badań. Zaangażowanie jest więc zjawiskiem dość ulotnym, które może w pełni przeminąć po wyciągnięciu grającego ze sfery skupienia[13].

Problemy pojawiają się również przy próbach retrospektwnego opisu i oceny zaangażowania po zakończeniu sesji gry. Brakuje w tym obszarze nauki pewnego wspólnego słownika, dzięki któremu uczestnicy badań mogliby w sposób jednoznaczny zwerbalizować subtelności tego złożonego doświadczenia[13]. Ograniczeni jesteśmy do używania bardzo ogólnych terminów takich jak "fajada", "zaangażowanie" czy "zaabsorbowanie", które nie oddają do końca bogactwa i głębi przeżywanych stanów[13].

Co więcej, zaangażowanie jest zjawiskiem silnie zależnym od otaczającego ją kontekstu, na który składać się może sama rozgrywka ale i również interakcje społeczne czy fizyczne środowisko, w którym znajduje się grający. Wszelkie próby jej wyizolowania i uproszczenia do czysto indywidualnego doświadczenia gruntownie znieksztalcają jego naturę.

W badaniach nad zaangażowaniem istotne są prace, które pomogły zdefiniować, gdzie ono rezyduje w odniesieniu do szerokiego spektrum wymiarów lub stanów poznawczych i emocjonalnych, takich jak motywacja i poczucie własnej skuteczności. Teoria samookreślenia pomaga wyjaśnić między innymi w jaki sposób motywacja wewnętrzna - napędzana czynnikami takimi jak potrzeba kompetencji - jest ważnym prekursorem zaangażowania. Zaangażowanie jest jednak czymś zgoła odmiennym od motywacji[33]. Można je postrzegać jako serię czasowych interakcji podczas wykonywania zadania, podczas gdy motywacja to bardziej uniwersalne osobiste nastawienie wobec nauki/zadania. Połączenie tych dwóch stanów, czyli zaangażowanie osoby w zadanie, do którego jest ona zmotywowana, może skutkować poczuciem satysfakcji i chęci ponownego zaangażowania się w to zadanie[33].

Przedstawione zostaną najczęściej występujące w literaturze terminy pokrewne lub stanowiące część zaangażowania:

- Imersja (ang. *Immersion*)
- Obecność (ang. *Presence*)
- Przepływ / trans (ang. *Flow*)
- Absorpcja (ang. *Absorption*)
- Dysocjacja (ang. *Dissociation*)

Imersja (zanurzenie) w grach wideo jest szeroko dyskutowanym pojęciem w pracach naukowych. Imersja zazwyczaj opisuje doświadczenie zaangażowania się w grę, przy zachowaniu pewnej świadomości otoczenia, a definiuje się ją także jako zdolność gry do wywoływanego uczucia bycia częścią, czy "obecności" w wirtualnym środowisku gry. Przewiduje się, że niemal każdy przeciętny grający doświadcza pewnego stopnia imersji[7]. Imersja (również nazywana "imersją sensoryczną i wyobrażeniową") oznaczać może również jak silne połączenie z grą odczuwał gracz[16]. Istnieją dyskusje na temat prawidłowej specyfikacji konstruktu imersji - czy powinien być modelowany jako refleksyjno-refleksyjny, czy refleksyjno-formatywny[3].

Obecność w kontekście gier wideo i wirtualnej rzeczywistości jest pojęciem, które wciąż ewoluje i oczekuje na ostateczną definicję, jednak zazwyczaj jest określana jako połączenie tych dwóch aspektów: 1. Bycie w normalnym stanie świadomości oraz 2. Doświadczanie poczucia znajdowania się wewnątrz wirtualnego środowiska[7]. Większość, ale nie wszyscy gracze, prawdopodobnie mają zdolność do doświadczania obecności w odpowiednich warunkach[7]. Poczucie obecności w alternatywnym środowisku może wynikać ze stymulacji sensorycznej[8]. Ogólniej, obecność wiąże się z emocjonalnym zaangażowaniem w grę[16].

Przepływ / trans najczęściej występuje w literaturze anglojęzycznej pod pojęciem *flow*. Przepływ można postrzegać jako głębokie, imersywne doświadczenie, które wynika z zaangażowania się osoby w zadanie o odpowiedniej równowadze pomiędzy wyzwaniem a poziomem umiejętności użytkownika[33]. Przepływ i rozgrywka w grach

są często łączone w kontekstach, gdzie użytkownik napotyka znaną (przykładowo z innych tytułów) strukturę rozgrywki natomiast doświadcza nowej treści (fabularnej, audio-wizualnej itd.)[33]. Przepływ to termin opisujący uczucie przyjemności, które występuje, gdy osiąga się równowagę pomiędzy umiejętnościami a wyzwaniem w procesie wykonywania wewnętrznie nagradzającej aktywności[7]. Posiadanie określonego celu i natychmiastowej informacji zwrotnej o wynikach zwiększa prawdopodobieństwo wystąpienia przepływu, a bycie w stanie przepływu wydaje się mieć wpływ na przebieg nauki użytkownika. Stany przepływu obejmują również uczucie kontroli, bycia jednością z aktywnością i doświadczanie zniekształceń czasu. Ponieważ wiąże się z doświadczaniem zmienionego stanu świadomości, doświadczenie przepływu może być nieco mniej powszechnie niż imersja czy obecność[7]. Z perspektywy modeli mentalnych przepływ do narracji występuje wtedy, gdy odbiorca całkowicie skupia się na czynności zrozumienia - tworzenia i aktualizowania modeli mentalnych reprezentujących historię[8].

Absorpcja to termin używany do opisania całkowitego zaangażowania w bieżące doświadczenie. W przeciwieństwie do imersji i obecności, a podobnie do przepływu, stan absorpcji psychologicznej wiąże się ze zmienionym stanem świadomości[7]. Ogólną skłonność osoby do wchodzenia w stan absorpcji psychologicznej można określić jako cechę, podczas gdy doświadczenie wchodzenia w ten stan w określonej aktywności najlepiej postrzegać jako stan przejściowy[7]. Niektóre badania dotyczące modeli zaangażowania w grę twierdzą, że absorpcja jest zasadniczo formą przepływu. Wtedy też lepiej wykluczyć ją z ostatecznego modelu zaangażowania w grę[3].

Dysocjacja znana jest jako objaw kliniczny występujący u osób cierpiących na traumę, natomiast naturalnie występuje również jej niepatologiczna forma[7]. Najbardziej poważnym przykładem dysocjacji niepatologicznej u dorosłych jest "hipnoza drogowa". Kierowcy wchodzą wtedy w stan absorpcji z aktywnością niezwiązaną z prowadzeniem. Równocześnie, konieczne obowiązki związane z manewrowaniem pojazdem są dalej wykonywane, pomimo że procesy mentalne związane z prowadzeniem są oddzielone od świadomej myśli[7].

3.2. SPOSOBY POMIARU ZAANGAŻOWANIA GRACZA

Istnieje wiele różnych podejść do pomiaru zaangażowania gracza, które cechują się odpowiednimi wadami i zaletami. Metody fizjologiczne, takie jak monitorowanie tężna, częstotliwości oddychania, aktywności mięśniowej (elektromiografia), fal mózgowych (elektroencefalografia) oraz przewodnictwo skórne, zapewniają większą obiektywność danych, ale są zwykle droższe (pod względem czasowym i finansowym) oraz trudniejsze w interpretacji[16]. Analiza zachowań graczy w grze (telemetria) również oferuje obiektywność, natomiast wciąż pewnego rodzaju wyzwanie stanowi redukcja złożoności związanej z profilowaniem

oraz powiązanie zachowań w grze z subiektywnym doświadczeniem gracza. Z drugiej strony, bardziej subiektywne metody oceny zaangażowania gracza, takie jak wywiady, grupy fokusowe, sondy w grze oraz kwestionariusze, są stosunkowo tańszymi alternatywami i charakteryzują się mniejszymi trudnościami w interpretacji niż pomiary fizjologiczne czy telemetria[16]. Wywiady, grupy fokusowe i sondy w grze pozwalają dogłębniej poznać odczucia grających, ale trudno je przeprowadzić na dużą skalę. W przeciwieństwie do nich, kwestionariusze można łatwo dystrybuować wśród bardzo dużych grup, i mimo że dostarczają mniej dogłębnych informacji niż inne subiektywne metody, umożliwiają skoncentrowanie się na konkretnych aspektach zaangażowania gracza[16].

W celu zewaluowania zaangażowania graczy w eksperymencie przeprowadzonym w ramach tejże pracy zdecydowano się na pomiar w formie kwestionariusza. Wybór konkretnego narzędzia ułatwiła praca Normana[24], który porównał ze sobą dwa modele kwestionariusza zaangażowania w grę (Game Engagement/Experience Questionnaire - GEQ)[7][12]. Ostatecznie wybrano model zaproponowany przez Brockmyera i współpracowników[7], który charakteryzuje się kilkoma istotnymi cechami. Po pierwsze, podejście Brockmyera i in.[7] koncentruje się na ocenie skłonności pojedynczego grającego do zaangażowania się w grę wideo, a nie na ocenie samej gry[24]. Autorzy[7] zaczęli od zaledwie 10 pozycji z pięciostopniową skalą ocen. Dodatkowe pozycje zostały stworzone, aby lepiej odzwierciedlić zaangażowanie, a cały kwestionariusz został następnie poddany ocenie na dwóch różnych próbach: 213 uczniów szkoły średniej i 51 studentów. Analizy wykazały potrzebę dodania kolejnych pozycji, doprowadzając do powstania 19-punktowego kwestionariusza. Mimo że nie wyjaśniono, skąd pochodziły nowe pozycje, kwestionariusz ten wykazuje dobre statystyki niezawodności[24], z współczynnikiem alfa Cronbacha na poziomie 0,85 oraz estymacją niezawodności osoby na poziomie 0,83 i niezawodności pozycji na poziomie 0,96 w modelu Rascha[7]. Ponadto, Brockmyer i współpracownicy[7] skupili się na jednowymiarowym kontinuum określonym jako "zaangażowanie", co jest główną zaletą ich pracy[24]. Autorzy przyznają jednak, że rekrutowali często grające osoby płci męskiej, aby zwiększyć prawdopodobieństwo doświadczenia głębskiego zaangażowania w grę w sytuacji laboratoryjnej, co sugerować może potrzebę ponownej ewaluacji kwestionariusza na szerszej próbce[24].

W ramach eksperimentu z kwestionariusza GEQ zaproponowanego przez Brockmyera i in.[7], cztery pozycje zostały zastąpione odpowiednikami pochodzącyymi z modelu IJsselsteijn'a[12]. Decyzja ta podyktywana była faktem, iż konkretne pozycje z kwestionariusza IJsselsteijn'a[12] lepiej odzwierciedlają nastawienie gracza wobec narracji przedstawianej w grze podczas pojedynczej sesji rozgrywki. Ponieważ eksperiment koncentruje się na ocenie zaangażowania w trakcie krótkich pojedynczych sesji, zastąpienie wybranych pozycji umożliwiło lepsze dopasowanie narzędzia do specyficznych warunków badania.

Number	Question
1	I lose track of time
2	I was interested in the game's story
3	I feel different
4	I felt that I could explore things
5	The game feels real
6	I was fully occupied with the game
7	I get wound up
8	Time seems to kind of stand still or stop
9	I feel spaced out
10	I was deeply concentrated in the game
11	I got tired ¹
12	Playing seems automatic
13	My thoughts go fast
14	I enjoyed it
15	I play without thinking how to play
16	Playing makes me feel calm
17	I play longer than I meant to
18	I really get into the game
19	I feel like I just can't stop playing

Tabela 3.1: Zmodyfikowany kwestionariusz GEQ

W tabeli 3.1 przedstawiono zmodyfikowany kwestionariusz Brockmyer'a[7] o cztery pozycje pochodzące bezpośrednio z modelu IJsselsteijn'a[12] (zaznaczone zostały poprzez pogrubienie).

¹ Oryginalnie w formie "I can't tell that I'm getting tired" co mogło być mylące w przypadku odpowiedzi tak/nie

4. PLANOWANY EKSPERYMENT

W tej sekcji przedstawiony zostanie pełny plan eksperymentu, w tym: projekt gry przeglądarkowej z gatunku *visual novel*, opis implementacji i wykorzystania generatywnych agentów opartych na dużych modelach językowych oraz sam przebieg badań.

4.1. PROJEKT GRY WYKORZYSTANEJ W EKSPERYMENCIE

Celem gry ma być oczywiście zbadanie wpływu wykorzystania generatywnych agentów na zaangażowanie grającego w narrację. W związku z tym rozgrywka powinna kłaść nacisk przede wszystkim na przedstawianie treści fabularnych, pozostawiając walory estetyczne czy mechaniki gry na drugim planie. Dodatkowo, całe doświadczenie powinno mieścić się w przedziale 15-30min i być dostępne w łatwy sposób (bez instalacji). Z uwagi na te założenia zdecydowano się na utworzenie gry dwuwymiarowej z gatunku *visual novel* (Patrz sekcja poniżej).

Gatunek *visual novel*

Forma *visual novel* (z ang. *powieść wizualna*) jest najczęściej uznawana jako gatunek gier komputerowych, choć niektórzy dostrzegają w niej zupełnie odrębne od gier medium[2].

Kluczowymi cechami *visual novel* są prezentacja tekstu za pomocą okienek dialogowych, które gracz musi kliknąć, aby przejść dalej, oraz statyczne grafiki przedstawiające postacie i otoczenie. Choć powieści wizualne często zawierają elementy multimedialne, takie jak animacje, muzyka czy dubbing, to nie są one ich kluczowymi składnikami[2].

Visual novel koncentrują się przede wszystkim na prezentacji narracji, z niewielką lub zerową ilością rozgrywki. Wiele z nich oferuje nieliniową, rozgałęzającą się fabułę z wieleoma zakończeniami i systemem wyborów wpływającym na dalszy przebieg wydarzeń[2]. Z drugiej strony, istnieją też powieści wizualne pozabawione jakiekolwiek rozgrywki i rozgałęzień fabularnych, określane mianem *kinetic novel*[1].

Ogólnie powieści wizualne wyróżniają się dominacją narracji przedstawianej za pomocą tekstu i grafik nad rozgrywką. Kryterium odróżniające je od gier przygodowych jest stopień, w jakim faktycznie wykorzystują mechaniki gry i gameplay w stosunku do narracji[2].

Projekt gry

Gra została opracowana przy wykorzystaniu silnika Ren'Py w wersji 8.2.1, opartego na języku Python. Ren'Py to popularne narzędzie służące do tworzenia gier tego rodzaju, oferujące zaawansowane możliwości pisania scenariuszy, zarządzania obrazami i dźwiękiem oraz tworzenia systemów wyborów i rozgałęzień fabularnych.

Wszystkie niezbędne zasoby graficzne pozyskane zostały ze społeczności twórców niezależnych na platformie itch.io. Grafiki postaci i tła zostały wyprodukowane przez twórców LinXueLian (<https://linxuelian.itch.io/>) oraz Potat0Master (<https://potat0master.itch.io/>), specjalizujących się w tego typu materiałach na potrzeby *visual novel* i gier przygodowych.

Po zakończeniu prac, gotowy produkt został wyeksportowany do postaci pliku wykonywalnego kompatybilnego z przeglądarkami internetowymi. Takie rozwiązanie umożliwia graczom pobieranie i uruchamianie tytułu bezpośrednio ze stron www, bez konieczności instalowania dodatkowego oprogramowania.

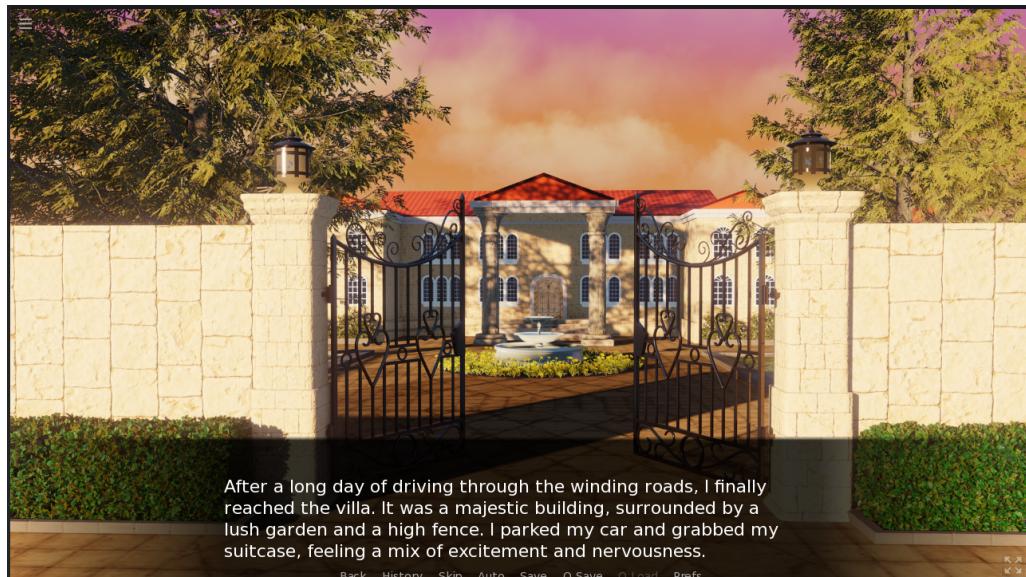
Finalną wersję gry opublikowano na platformie dystrybucji cyfrowej itch.io, która stała się głównym kanałem udostępniania produkcji odbiorcom. Itch.io jest popularnym hubem dla niezależnych twórców gier, w tym deweloperów *visual novel* wykorzystujących silniki takie jak Ren'Py.

Na rysunku 4.1 przedstawiono ekran początkowy widziany przez gracza po uruchomieniu gry. Panele ustawień, zarządzania zapisami / wczytaniami gry oraz pomocy są automatycznie generowane przez silnik Ren'Py.



Rys. 4.1: Ekran startowy gry

Przykład prezentacji treści fabularnej jest widoczny na rysunku 4.2. Widok zasadniczo składa się z tła, opcjonalnie z postaci widocznej na pierwszym planie oraz z panelu tekstu-wego (W tym przypadku nie widać imienia postaci więc grający powinien dany fragment uznać za monolog wewnętrzny głównego bohatera).



Rys. 4.2: Wprowadzenie fabularne / przykład narracji

Przykładowy widok dialogu z jedną z postaci NPC został przedstawiony na rysunku 4.3.



Rys. 4.3: Przykładowy dialog z postacią NPC

Zarys fabularny

Gracz zostaje zaproszony na imprezę organizowaną przez Sylvię, bogatą i ekstrawaganczką kobietę. Na przyjęciu pojawiają się również: Adam - marzyciel pracujący w piekarni ojca, Nathaniel - bezwzględny biznesmen, Randy - genialny pianista o ciemnej stronie, Mary - zmagająca się z problemami finansowymi pisarka kryminałów oraz Florian - bogaty prawnik zakochany w Sylvii. Podczas imprezy Sylvia zostaje zamordowana. Gracz znajduje jej ciało i zostawiony list, który sugeruje jej samobójstwo po czym wzywa policję. Do przyjazdu służb potrzeba kilku godzin, więc główny bohater postanawia samemu rozwiązać tę sprawę.

Gracz musi rozmawiać z pozostałymi postaciami NPC, zbierając od nich informacje na temat ich relacji z Sylvią, motywów i alibi w nocy morderstwa. Adam był skrycie zakochany w Sylvii i zazdrosny o jej związek z Florianem. Nathaniel próbował bezskutecznie przejąć piekarnię ojca Adama. Randy to utalentowany, ale arogancki pianista z ciemną przeszłością i problemami z prawem, skrywający wiele tajemnic. Mary to pisarka kryminałów zmagająca się z problemami finansowymi, będąca najlepszą przyjaciółką Sylvii. Florian to bogaty prawnik zakochany w Sylvii, będący rywalem Adama w miłości, choć nie zdawał sobie z tego sprawy. Każda postać ma własny charakter i sekrety, które mogą okazać się istotne dla śledztwa.

Gra kończy się konfrontacją, podczas której gracz wskazuje sprawcę. W zależności od ostatecznego wyboru może być przedstawione "dobre" lub "złe" zakończenie.

4.2. OPIS GENERATYWNYCH AGENTÓW

Inworld AI to innowacyjna platforma umożliwiająca tworzenie fotorealistycznych, interaktywnych postaci wirtualnych, które mogą być wykorzystywane w różnorodnych zastosowaniach. Firma ta oferuje bezkodową technologię, dzięki której projektowanie postaci staje się niezwykle łatwe i może zostać przeprowadzone w ciągu zaledwie kilku minut.

Zakres możliwych zastosowań postaci stworzonych za pomocą Inworld AI jest szeroki i obejmuje m.in. gry wideo otwartego świata, wirtualne awatary i ambasadorów marek, immersywne doświadczenia edukacyjne oraz szkolenia. Platforma ta umożliwia tworzenie realistycznie wyglądających i zachowujących się postaci niezależnych (NPC), które mogą wchodzić w interakcje z użytkownikami w sposób zbliżony do interakcji z prawdziwymi ludźmi.

W miarę jak przestrzeń sztucznej inteligencji ewoluje w bardzo szybkim tempie, Inworld AI stale aktualizuje swoje funkcjonalności, aby optymalizować proces tworzenia postaci i utrzymać pozycję lidera w dziedzinie generatywnej sztucznej inteligencji. Firma ta kładzie nacisk na ciągłe doskonalenie swoich technologii, aby zapewnić użytkownikom możliwość tworzenia coraz bardziej realistycznych i zaawansowanych postaci wirtualnych.

Zastosowanie postaci wygenerowanych przez Inworld AI może przynieść korzyści w wielu branżach, takich jak rozrywka, marketing, edukacja czy szkolenia. Dzięki ich realistycznemu wyglądowi i zachowaniu, mogą one zwiększać zaangażowanie użytkowników oraz oferować bardziej immersyjne i angażujące doświadczenia.

Kluczowe funkcjonalności platformy Inworld AI

Wspólna Wiedza (Common Knowledge) umożliwia zdefiniowanie ogólnej wiedzy, która ma być znana przez wiele postaci. Może to obejmować informacje światobudujące, które wszystkie postacie powinny znać lub wiedzę przypisaną tylko do konkretnej grupy postaci, np. tych, które wiedzą, że w ich świecie występuje magia i potrafią jej używać.

Podstawowy Opis (Core Description) stanowi fundament osobowości postaci i w znacznym stopniu wpływa na wszystkie jej późniejsze reakcje. Powinien koncentrować się na szczegółach dotyczących obecnych okoliczności życiowych postaci, jej historii, nadziei oraz sposobu w jaki się prezentuje. W tej sekcji można również wspomnieć o kluczowych relacjach, przedsięwzięciach biznesowych czy lokalizacjach związanych z daną postacią. Jeśli postać posługuje się specyficzną manierą mówienia lub przestrzega określonych reguł zachowania, można to tutaj dodać.

W sekcji Motywacje (Motivations) należy umieścić pojedyncze zdania opisujące co motywuje postać do rozmów z innymi. Może to być chęć realizacji celu lub pragnienia, przedstawienia swojej opinii lub pomocy użytkownikowi w zdobyciu wiedzy. Ważne jest, aby zastanowić się, co napędza daną postać, ponieważ będzie to wpływać na jej reakcje, a ona sama będzie poszukiwać okazji do wplatania swoich motywacji w rozmowę.

W polu Wady (Flaws) wpisuje się pojedyncze zdania dotyczące niedoskonałości i lęków postaci. Co powstrzymuje ją przed realizacją motywacji? W jaki sposób wyraża swoje wewnętrzne konflikty poprzez zewnętrzny dialog? Jakie tematy wywołują negatywną reakcję?

Rola (Role) zapewnia ramy określające, w jaki sposób postać wchodzi w interakcje z otaczającym ją światem. Może to być coś ogólnego jak "Bohater" lub "Asystent", ale bardziej szczegółowe archetypy lub zawody, takie jak "Średniowieczny Wojownik" czy "Ambasador Marki X", lepiej ukierunkującą sztuczną inteligencję Inworld.

Zainteresowania i Hobby (Hobbies and Interests) to krótka lista hobby i zainteresowań postaci. Może się do nich odwoływać w rozmowie. Mogą być one szerokie (np. pomaganie w rozwiązywaniu problemów użytkowników) lub specyficzne dla motywacji postaci (np. zasadzanie się na rywalizujące gangi).

Cechy Osobowości (Personality Traits) pomogą stworzyć odpowiednią osobowość i reakcje sztucznej inteligencji Inworld.

Nastrój i Suwerki Osobowości (Mood and Personality Sliders) decydują o rodzajach emocji, które będzie przejawiać postać w odpowiedzi na interakcje. Wpływają również

na ton jej wypowiedzi - jeśli stworzymy radosną postać, zazwyczaj będzie odpowiadać w radosny sposób.

Fakty i Wiedza (Facts and Knowledge) pomagają w predefiniowaniu odpowiedzi na pytania użytkowników. Wiedza Osobista to wszystko co postać wie osobiście, natomiast Wspólna Wiedza to miejsce, gdzie można dodać szersze informacje np. o epoce lub świecie gry, które mogą być współdzielone między wieloma postaciami. Filtry Wiedzy (Knowledge Filters) mają na celu redukcję odstępstw i niespójności, które mogłyby wykroczać poza ustalone parametry postaci. Oferowane są trzy poziomy filtrów: Ścisły, Łagodny i Brak Filtra.

Sekcja Cele (Goals) umożliwia ustawienie specyficznych wyzwalaczy, które spowodują, że postać zareaguje w określony sposób w konkretnych scenariuszach i interakcjach. Cel działa jako "mechanizm konsekwencji", który jest aktywowany przez zdarzenie wyzwalające i inicjuje określoną akcję. Pozwala to twórcom na lepszą kontrolę nad postaciami w czasie rzeczywistym. Dodatkowo, system celów monitoruje ich osiąganie, wysyłając wyraźny sygnał do klienta po zakończeniu celu.

Styl Dialogu (Dialogue Style) umożliwia wybór spośród szeregu predefiniowanych stylów lub stworzenie własnego, niestandardowego stylu. Ta funkcja, w połączeniu z Osobowością i Emocjami, zdecyduje o sposobie, w jaki postać będzie prezentować swoje odpowiedzi. Może być ciekawa i zadawać mnóstwo pytań lub być tajemnicza i niewiele zdradzać.

Mutacje Postaci (Character Mutations) mają na celu wprowadzanie tymczasowych zmian w atrybutach postaci, dając więcej kontroli nad nią. Funkcja ta umożliwia implementację postaci przechodzących przemijające zmiany. Mutacje są częścią systemu celów i pozwalają na modyfikację aktualnie aktywnej postaci jako akcję celu.

Sceny (Scenes) dostarczają kontekstu, opisując bezpośrednio otoczenie postaci. Jeśli planujemy po prostu rozmawiać z postacią w studio Inworld lub w Inworld Arcade, nie będziemy musieli dodawać sceny, chyba że jest to potrzebne dla dodatkowego kontekstu.

Opis wykrojonych postaci

Adam jest marzycielem z małego miasteczka, który pracuje w piekarni swojego ojca. Marzy o lepszym życiu i pragnie wyrwać się z ciasnych ram swojej codzienności. Zna Floriana z liceum i był potajemnie zakochany w Sylvii. Czuł zazdrość wobec Floriana z powodu jego relacji z Sylvią. Adam nienawidzi Nathaniela, który chce wykupić piekarnię jego ojca. Nie przepada również za Randym, którego uważa za aroganta. W nocy, gdy doszło do morderstwa, był przebudzony, obserwując Sylvię, i to on znalazł jej ciało.

Florian pochodzi z bogatej rodziny prawniczej. Choć miał wszystko, brakowało mu prawdziwej miłości, którą znalazł dopiero w Sylvii. Jest osobą lojalną i gotową do poświęceń dla bliskich. Przyjaźnił się z Adamem od czasów liceum, choć nie wiedział o jego uczuciach do Sylvii. Wiedział, że Sylvia była najlepszą przyjaciółką Mary. Florian nie lubi Nathaniela

za jego metody działania, ale podziwia talent Randy'ego, nie wiedząc o jego problemach z prawem. Twierdzi, że spał przez całą noc i nie widział Sylvii przed jej śmiercią.

Mary jest pisarką powieści kryminalnych, która zmaga się z finansowymi problemami. Jej najlepszą przyjaciółką była Sylvia, która zawsze ją wspierała. Znała Sylvię bardzo dobrze, wiedziała o jej depresji i problemach. Była w dobrych stosunkach z Nathaniel'em, który próbował pomóc jej w znalezieniu wydawcy. Mary nie ufa Randy'emu i uważa go za podejrzanego. Gdy doszło do morderstwa, była w łazience i usłyszała krzyk, ale nie widziała, kto zabił Sylvię.

Nathaniel jest biznesmenem, który dorobił się majątku dzięki swojej determinacji i bezwzględności. Pragnie zdobyć szacunek, którego nigdy nie zaznał. Podobał mu się Adam, choć ich relacja była napięta z powodu interesów. Nathaniel cenił Randy'ego za jego talent i wierzył, że po odbyciu kary więzienia, Randy stał się lepszą osobą. Myślał, że Mary mogłaby zabić z desperacji. Twierdzi, że spędził noc w kuchni, widząc Randy'ego palącego papierosy na zewnątrz.

Randy to utalentowany pianista z trudną przeszłością. Choć jest podziwiany za swój talent, jego charakter pozostawia wiele do życzenia – jest arogancki i porywczy. Uważał, że Adam jest podejrzany z powodu jego zachowania względem Sylvii. Miał mieszane uczucia co do Floriana, którego uważa za podejrzanego. Wierzył, że Mary potrzebuje prawdziwej inspiracji do swojej pracy. Twierdził, że spędził noc na zewnątrz, paląc papierosy. W rzeczywistości to on zabił Sylvię, choć stara się to ukryć.

Wykorzystanie agentów w praktyce

lorfloefleofe

4.3. ZAPLANOWANY PRZEBIEG EKSPERYMENTU

pofkepfokawpefkwpeofk

Forma eksperymentu

lorfloefleofe

Gromadzone dane

lorfloefleofe

5. WYNIKI

fuwenuiwhfe

PODSUMOWANIE

Curabitur tellus magna, porttitor a, commodo a, commodo in, tortor. Donec interdum. Praesent scelerisque. Maecenas posuere sodales odio. Vivamus metus lacus, varius quis, imperdiet quis, rhoncus a, turpis. Etiam ligula arcu, elementum a, venenatis quis, sollicitudin sed, metus. Donec nunc pede, tincidunt in, venenatis vitae, faucibus vel, nibh. Pellentesque wisi. Nullam malesuada. Morbi ut tellus ut pede tincidunt porta. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam congue neque id dolor.

SPIS RYSUNKÓW

1.1	Spacewar (1962)	8
1.2	Pong (1972)	8
1.3	Breakout (1976)	9
1.4	Space Invaders (1978)	10
1.5	Super Mario Bros. (1985)	10
1.6	The Legend of Zelda (1986)	11
1.7	Crash Bandicoot (1996)	12
1.8	Half-life (1998)	13
1.9	Life is Strange (2015)	14
1.10	Wiedźmin 3 (2015)	15
1.11	Final Fantasy I (1987)	16
1.12	Final Fantasy VII (1997)	17
1.13	Final Fantasy X (2001)	17
1.14	Final Fantasy XIII (2009)	18
1.15	Liniowa struktura gry[15]	19
1.16	Struktura łańcucha pereł	20
1.17	Struktura rozgałęzająca się	21
1.18	The Dark Pictures Anthology: Man of Medan (2019) - Supermassive Games	23
1.19	Bloki tekstu służące do przedstawienia fabuły	23
1.20	Dialog z NPC - "Wiedźmin 3" (2015)	24
1.21	Final Fantasy VII (1997)	27
1.22	Life is Strange (2015)	27
1.23	Fallout 4 (2015) + wersja zmodyfikowana[10]	28
1.24	Fragment spisu ikon dialogowych - "Dragon Age: Inquisition" (2014)[9]	29
1.25	Przykład dialogu z wykorzystaniem ikony emocji - "Dragon Age: Inquisition" (2014)	29
1.26	Fallout: New Vegas (2010)	30
1.27	Wiedźmin 3 (2015)	31
1.28	Ogólny wygląd rozgrywki w "Otochłani"	32
1.29	Dialog w "Otochłani"	32
1.30	Polecenia dostępne w "Otochłani"	33
1.31	Fragment samouczka "Otochłani"	33
2.1	Fabuła "Czerwonego Kapturka" zapisana za pomocą POP	35
2.2	Przykładowy proces decyzyjny Markowa	37
4.1	Ekran startowy gry	49

4.2 Wprowadzenie fabularne / przykład narracji	50
4.3 Przykładowy dialog z postacią NPC	50

SPIS LISTINGÓW

SPIS TABEL

1.1	Dwanaście etapów wzorca narracyjnego "Podróży bohatera" [15]	7
3.1	Zmodyfikowany kwestionariusz GEQ	47

Dodatki

A. SŁOWNIK

Nulla ac nisl. Nullam urna nulla, ullamcorper in, interdum sit amet, gravida ut, risus. Aenean ac enim. In luctus. Phasellus eu quam vitae turpis viverra pellentesque. Duis feugiat felis ut enim. Phasellus pharetra, sem id porttitor sodales, magna nunc aliquet nibh, nec blandit nisl mauris at pede. Suspendisse risus risus, lobortis eget, semper at, imperdiet sit amet, quam. Quisque scelerisque dapibus nibh. Nam enim. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nunc ut metus. Ut metus justo, auctor at, ultrices eu, sagittis ut, purus. Aliquam aliquam.

B. INSTRUKCJE INWORLD.AI DLA POSTACI

Nulla ac nisl. Nullam urna nulla, ullamcorper in, interdum sit amet, gravida ut, risus. Aenean ac enim. In luctus. Phasellus eu quam vitae turpis viverra pellentesque. Duis feugiat felis ut enim. Phasellus pharetra, sem id porttitor sodales, magna nunc aliquet nibh, nec blandit nisl mauris at pede. Suspendisse risus risus, lobortis eget, semper at, imperdiet sit amet, quam. Quisque scelerisque dapibus nibh. Nam enim. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nunc ut metus. Ut metus justo, auctor at, ultrices eu, sagittis ut, purus. Aliquam aliquam.

BIBLIOGRAFIA

- [1] *Kinetic novel*, <https://tvtropes.org/pmwiki/pmwiki.php/Main/KineticNovel>. Data dost. 18 marca 2024.
- [2] *Visual novel*, <https://tvtropes.org/pmwiki/pmwiki.php/Main/VisualNovel>. Data dost. 18 marca 2024.
- [3] Abbasi, A.Z., Ting, D.H., Hlavacs, H., *Engagement in games: Developing an instrument to measure consumer videogame engagement and its validation*, International Journal of Computer Games Technology. 2017, tom 2017, str. 7363925.
- [4] Al-Nassar, S., Schaap, A., Zwart, M.V.D., Preuss, M., Gómez-Maureira, M.A., *Questville: Procedural quest generation using nlp models*, w: *Proceedings of the 18th International Conference on the Foundations of Digital Games*, FDG '23 (Association for Computing Machinery, New York, NY, USA, 2023).
- [5] Ammanabrolu, P., Broniec, W., Mueller, A., Paul, J., Riedl, M.O., *Toward automated quest generation in text-adventure games*. 2020.
- [6] Assaf, M., *From pong to narrative: The evolution of ai in gaming*. 2023, tom 108.
- [7] Brockmyer, J.H., Fox, C.M., Curtiss, K.A., McBroom, E., Burkhart, K.M., Pidruzny, J.N., *The development of the game engagement questionnaire: A measure of engagement in video game-playing*, Journal of Experimental Social Psychology. 2009, tom 45, 4, str. 624–634.
- [8] Busselle, R., Bilandzic, H., *Measuring narrative engagement*, Media Psychology. 2009, tom 12, 4, str. 321–347.
- [9] *Dialogue wheel (inquisition)*, [https://dragonage.fandom.com/wiki/Dialogue_wheel_\(Inquisition\)](https://dragonage.fandom.com/wiki/Dialogue_wheel_(Inquisition)). Data dost. 10 kwietnia 2024.
- [10] Fraser, J., Papaioannou, I., Lemon, O., *Spoken conversational ai in video games: Emotional dialogue management increases user engagement*, w: *Proceedings of the 18th International Conference on Intelligent Virtual Agents*, IVA '18 (Association for Computing Machinery, New York, NY, USA, 2018), str. 179–184.
- [11] Hausknecht, M., Ammanabrolu, P., Côté, M.A., Yuan, X., *Interactive fiction games: A colossal adventure*. 2020.
- [12] IJsselsteijn, W., de Kort, Y., Poels, K., *The Game Experience Questionnaire* (Technische Universiteit Eindhoven, 2013).
- [13] Ijsselsteijn, W., De Kort, Y., Poels, K., Jurgelionis, A., Bellotti, F., *Characterising and measuring user experiences in digital games*, Journal of Applied Mechanics-transactions of The Asme - J APPL MECH. 2007.
- [14] *Interactive fiction*, https://en.wikipedia.org/wiki/Interactive_fiction. Data dost. 9 kwietnia 2024.

- [15] Ip, B., *Narrative structures in computer and video games: Part 1: Context, definitions, and initial findings*, Games and Culture. 2011, tom 6, 2, str. 103–134.
- [16] Johnson, D., Gardner, M.J., Perry, R., *Validation of two game experience scales: The player experience of need satisfaction (pens) and game experience questionnaire (geq)*, International Journal of Human-Computer Studies. 2018, tom 118, str. 38–46.
- [17] Kryah, K., *The evolution of final fantasy's narratives*, <https://the-artifice.com/final-fantasy-evolution-narratives/>. Data dost. 18 marca 2024.
- [18] Madsen, H., *25 years later, the most overlooked final fantasy deserves more credit*, <https://www.inverse.com/gaming/25-years-later-the-most-overlooked-final-fantasy-deserves-more-credit>. Data dost. 18 marca 2024.
- [19] Magerko, B., Laird, J., Assanie, M., Kerfoot, A., Stokes, D., *Ai characters and directors for interactive computer games*. 2004.
- [20] Majewski, J., i in., *Theorising video game narrative*, Bond University. 2003.
- [21] Markowicz, W., *Otchłań - historia zapomnianej gry tekstowej*, <https://www.youtube.com/watch?v=zP1twKbLmAw>. Data dost. 18 marca 2024.
- [22] Mateas, M., Stern, A., *A behavior language for story-based believable agents*, IEEE Intelligent Systems. 2002, tom 17, 4, str. 39–47.
- [23] Montfort, N., *Interactive fiction's fourth era*, https://nickm.com/if/fourth_era.html. 2008. Data dost. 15 kwietnia 2024.
- [24] Norman, K.L., *Geq (game engagement/experience questionnaire): A review of two papers*, Interacting with Computers. 2013, tom 25, 4, str. 278–283.
- [25] OpenAI, Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F.L., Almeida, D., Altenschmidt, J., i in., *Gpt-4 technical report*. 2024.
- [26] Papworth, S., *Storytelling through Gameplay: Dimensions of AI Design for Narrative Purposes*, Rozprawa doktorska. 2016.
- [27] Park, J.S., O'Brien, J.C., Cai, C.J., Morris, M.R., Liang, P., Bernstein, M.S., *Generative agents: Interactive simulacra of human behavior*. 2023.
- [28] Ramirez, A., Bulitko, V., *Automated planning and player modeling for interactive storytelling*, IEEE Transactions on Computational Intelligence and AI in Games. 2015, tom 7, 4, str. 375–386.
- [29] Riedl, M., Thue, D., Bulitko, V., *Game AI as Storytelling* (Springer New York, New York, NY, 2011), str. 125–150.
- [30] Stone, C., *The evolution of video games as a storytelling medium, and the role of narrative in modern games*, <https://www.gamedeveloper.com/design/the-evolution-of-video-games-as-a-storytelling-medium-and-the-role-of-narrative-in-modern-games>. Data dost. 18 marca 2024.
- [31] Umbraško, A., Drury, R., *Applying chatgpt in ai-based dynamic video game narrative generation system*. 2023.
- [32] Westera, W., Prada, R., Mascarenhas, S., Santos, P.A., Dias, J., Guimarães, M., Georgiadis, K., Nyamsuren, E., Bahreini, K., i in., *Artificial intelligence moving serious gaming: Presenting*

- reusable game ai components*, Education and Information Technologies. 2020, tom 25, 1, str. 351–380.
- [33] Wiebe, E.N., Lamb, A., Hardy, M., Sharek, D., *Measuring engagement in video game-based environments: Investigation of the user engagement scale*, Computers in Human Behavior. 2014, tom 32, str. 123–132.
- [34] Yan, M., Li, R., Zhang, H., Wang, H., Yang, Z., Yan, J., *Larp: Language-agent role play for open-world games*. 2023.