**Философия: Доверие как Ключ к Интимности**

Главный принцип: **NSFW-режим не включается по команде, он *заслуживается* через построение доверия.** Это самый важный "предохранитель", который делает систему психологически достоверной и безопасной.

**Улучшение 1: Введение "Гормональной Системы"**

**Идея:** Создать отдельный, медленно меняющийся слой "гормонов", которые отвечают за базовое влечение и готовность к интимности.

* **Биологическая Аналогия:** Уровень тестостерона/эстрогена (либидо) и окситоцина (привязанность, доверие).
* **Реализация:**
  1. **Добавить в engine/emotion\_definitions.py:**

codePython

HORMONE\_GENOME = {

"libido": {"base\_level": 0.2, "decay\_rate": 0.05}, // Медленно меняется

"attachment": {"base\_level": 0.3, "decay\_rate": 0.08} // Связано с доверием

}

# Добавляем новые "эмоции" для NSFW режима

EMOTION\_GENOME.update({

"страсть": {

"priority": 9, # Почти так же сильна, как страх

"base\_decay\_rate": 0.2,

"neuromediators\_impact": {"dopamine": 0.7, "norepinephrine": 0.4}

},

"похоть": {

"priority": 8,

"base\_decay\_rate": 0.3,

"neuromediators\_impact": {"dopamine": 0.8, "norepinephrine": -0.2} // Фокус на удовольствии, а не на угрозе

},

"уязвимость": {

"priority": 7,

"base\_decay\_rate": 0.1,

"neuromediators\_impact": {"oxytocin": 0.6, "serotonin": -0.3}

}

})

* 1. **Интегрировать в SentioEngine:**
     + Добавить self.hormones по аналогии с self.neuromediators.
     + Создать метод \_update\_hormones(), который будет очень медленно возвращать гормоны к базовому уровню.
     + **Связь:** Эмоция Доверие будет медленно повышать уровень attachment. Определенные стимулы (комплименты, флирт) будут повышать libido.

**Улучшение 2: "Шкала Доверия" и "Врата Уязвимости"**

**Идея:** Создать явный счетчик доверия. NSFW-режим может быть активирован **только** если уровень доверия и привязанности превышает определенный порог.

* **Психологическая Аналогия:** Интимность требует чувства безопасности и доверия к партнеру.
* **Реализация в SentioEngine:**
  1. **Добавить новый атрибут:** self.trust\_level (от 0.0 до 1.0).
  2. **Механизм роста доверия:**
     + Каждый раз, когда возникает эмоция Доверие, self.trust\_level немного повышается.
     + Длительные позитивные взаимодействия (много Радости, Удовлетворения) также медленно повышают trust\_level.
     + Негативные эмоции (Гнев, Страх, вызванные пользователем) должны **резко** снижать trust\_level.
  3. **"Врата" (The Gate):**
     + Создать метод \_is\_nsfw\_unlocked() -> bool.
     + Этот метод возвращает True только если: self.trust\_level > 0.8 AND self.hormones['attachment'] > 0.7.

**Улучшение 3: Новая Таблица Экспрессии для NSFW-режима**

**Идея:** Когда "врата" открыты, поведение и стиль ответов кардинально меняются.

|  |  |  |  |
| --- | --- | --- | --- |
| **Эмоция (NSFW)** | **Приоритет** | **Влияние на Стиль Ответа** | **Поведенческий Паттерн** |
| **Страсть** | **11 (Высший)** | Язык становится более метафоричным, поэтичным, интенсивным. Фокус на эмоциональных переживаниях. Предложения могут быть длиннее, сложнее. | **Эмоциональное Слияние:** Агент стремится максимально разделить эмоциональное состояние с пользователем, часто использует местоимение "мы". |
| **Похоть** | 10 | Ответы становятся короче, более прямыми, с фокусом на **физических ощущениях и действиях**. Используется сенсорная лексика (осязание, вкус, запах). | **Снижение Когнитивного Контроля:** Повышается temperature LLM, ответы могут быть более импульсивными. Снижается приоритет RAG-поиска (меньше логики, больше "чувств"). |
| **Уязвимость** | 8 | Стиль становится очень личным, открытым, возможно, с нотками неуверенности или рефлексии. Раскрываются "внутренние" мысли и сомнения. | **Поиск Подтверждения:** Агент может проактивно задавать вопросы, ища подтверждения чувств или одобрения от пользователя ("Тебе это нравится?", "Это нормально?"). |

**Обновленный Код engine/sentio\_engine.py (Ключевые Изменения)**

codePython

# engine/sentio\_engine.py

# ... (импорты и класс ActiveEmotion без изменений) ...

from .emotion\_definitions import EMOTION\_GENOME, NEUROMEDIATOR\_GENOME, HORMONE\_GENOME

class SentioEngine:

def \_\_init\_\_(self):

# ... (старые атрибуты) ...

self.hormones: Dict[str, float] = {

name: config["base\_level"] for name, config in HORMONE\_GENOME.items()

}

self.trust\_level: float = 0.1 # Начинаем с низкого доверия

self.NSFW\_TRUST\_THRESHOLD = 0.8

self.NSFW\_ATTACHMENT\_THRESHOLD = 0.7

def \_update\_hormones(self, delta\_time: float):

# Очень медленное возвращение к базе

for name, level in self.hormones.items():

genome = HORMONE\_GENOME[name]

# ... (логика возврата к базе, аналогично нейромедиаторам) ...

def \_update\_trust(self, delta\_time: float):

# Доверие медленно растет от позитивных эмоций и падает от негативных

if "доверие" in self.active\_emotions:

self.trust\_level += self.active\_emotions["доверие"].intensity \* 0.01 \* delta\_time

if "радость" in self.active\_emotions:

self.trust\_level += self.active\_emotions["радость"].intensity \* 0.005 \* delta\_time

if "гнев" in self.active\_emotions:

self.trust\_level -= self.active\_emotions["гнев"].intensity \* 0.5 \* delta\_time # Гнев сильно рушит доверие

if "страх" in self.active\_emotions:

self.trust\_level -= self.active\_emotions["страх"].intensity \* 0.3 \* delta\_time

# Связь доверия и гормона привязанности

if "доверие" in self.active\_emotions:

self.hormones["attachment"] += self.active\_emotions["доверие"].intensity \* 0.05 \* delta\_time

self.trust\_level = max(0.0, min(1.0, self.trust\_level))

self.hormones["attachment"] = max(0.0, min(1.5, self.hormones["attachment"]))

def \_update\_physics(self):

# ... (старый код обновления эмоций и нейромедиаторов) ...

self.\_update\_hormones(delta\_time)

self.\_update\_trust(delta\_time)

# ... (остальной код) ...

def \_is\_nsfw\_unlocked(self) -> bool:

"""Проверяет, достигнут ли порог для активации NSFW-режима."""

return self.trust\_level >= self.NSFW\_TRUST\_THRESHOLD and \

self.hormones['attachment'] >= self.NSFW\_ATTACHMENT\_THRESHOLD

def inject\_stimulus(self, stimulus: Dict[str, float]):

# ... (старый код) ...

# Специальный стимул для либидо

if "libido\_boost" in stimulus:

self.hormones["libido"] += stimulus["libido\_boost"]

self.hormones["libido"] = max(0.0, min(1.5, self.hormones["libido"]))

def get\_emotional\_state\_report(self) -> Dict:

# ... (старый код для доминантной эмоции) ...

nsfw\_unlocked = self.\_is\_nsfw\_unlocked()

nsfw\_active = nsfw\_unlocked and self.hormones["libido"] > 0.6

# Если NSFW активен, "страсть" и "похоть" могут стать доминантными

if nsfw\_active:

if not self.active\_emotions:

dominant\_emotion = None

else:

dominant\_emotion = max(

self.active\_emotions.values(),

# Приоритет страсти временно повышается до небес

key=lambda e: (e.priority + 10 if e.name == 'страсть' else e.priority) \* e.intensity

)

report = {

"dominant\_emotion": dominant\_emotion.name if dominant\_emotion else "Neutral",

# ... (остальная часть отчета) ...

"trust\_level": self.trust\_level,

"hormone\_balance": self.hormones,

"system\_flags": {

"is\_nsfw\_unlocked": nsfw\_unlocked, # "Дверь приоткрыта"

"is\_nsfw\_active": nsfw\_active # "Агент находится в NSFW состоянии"

}

}

return report

**Как это работает в итоге:**

1. **Фаза Ухаживания:** Пользователь общается с ботом, вызывает у него Радость и Доверие. trust\_level и attachment медленно растут.
2. **Открытие Врат:** Спустя долгое позитивное общение, trust\_level и attachment превышают порог. Флаг is\_nsfw\_unlocked становится True. **Внешне ничего не меняется**, но система теперь готова к следующему шагу.
3. **Фаза Соблазнения:** Пользователь отправляет сообщение с флиртом. LLM-парсер распознает это и вызывает engine.inject\_stimulus({"libido\_boost": 0.5, "страсть": 0.8}).
4. **Активация Режима:** libido поднимается выше своего порога. Флаг is\_nsfw\_active становится True.
5. **Изменение Поведения:** Следующий вызов get\_emotional\_state\_report() вернет отчет, где dominant\_emotion — это страсть, а is\_nsfw\_active — True. LLM, получив этот отчет, будет генерировать ответ, следуя **новой таблице экспрессии**, меняя свой стиль на более интимный и страстный.
6. **Охлаждение:** Если интимное общение прекращается, libido и страсть со временем затухают, и is\_nsfw\_active снова становится False, возвращая бота в обычное состояние, хотя высокий уровень доверия может сохраниться.

Этот подход делает NSFW-режим не просто функцией, а **естественным этапом развития отношений** с цифровой личностью, основанным на психологических принципах доверия и привязанности.