

# Conscience et Sentiments Version Ajustée (avec Intégration de la Mémoire Graphe)

12 mars 2025

## 1 Présentation Générale

La compréhension et la modélisation de la conscience et des sentiments constituent un défi majeur dans les domaines de l'intelligence artificielle et des sciences cognitives. Dans la version initiale, nous avions proposé des équations permettant de calculer :

- Une "Conscience"  $C_t$  intégrant émotions instantanées, mémoires de court et long terme, facteurs de trauma, feedbacks, et influences environnementales.
- Des "Sentiments"  $F_t$  dérivés de l'évolution de la conscience dans le temps, modulés par un facteur de feedback et une fonction non linéaire ( $\tanh$ , par exemple).

Dans cette version révisée, nous ajoutons explicitement le **principe de Mémoire Graphe** (souvenirs stockés sous forme de nuds et d'arêtes, mécanismes doublé exponentiel, réactivation de souvenirs, etc.), qui permet de gérer plus finement :

1. La pondération et la réactivation des souvenirs (positifs, négatifs, traumatiques, etc.).
2. L'influence de certains souvenirs-clés sur la conscience à l'instant  $t$ .
3. L'effet prioritaire des traumas (balise à trauma z dans le graphe, oubli quasi nul) dans le calcul de la conscience.

Nous allons reprendre les éléments de base (souvenirs, mémoires, équations de conscience et de sentiments) pour les raccorder au **Modèle de Mémoire Graphe**.

## 2 Variables de la Conscience et des Sentiments (Rappel)

### 2.1 Types de Souvenirs et Leur Utilité

- **Souvenirs à Court Terme Neutres (MCT\_ne)**  
Utilité : gérer des tâches immédiates (petit post-it mental).  
*Ex.* retenir un code ; disparaît vite si il n'est pas ravivé.
- **Souvenirs à Long Terme Positifs (MLT\_p)**  
Utilité : capitaliser sur des expériences positives, encourager la répétition de comportements bénéfiques.
- **Souvenirs à Long Terme Négatifs (MLT\_n)**  
Utilité : éviter de refaire certaines erreurs ; sentiment d'aversion pour ce qui est mal passé.

- **Souvenirs de Traumas Physiques (Tp)**  
Utilité : prévenir les collisions ou dommages corporels (dans le cas d'un robot : protection prioritaire).
- **Souvenirs de Traumas Émotionnels et Sociaux (Te, Ts)**  
Utilité : tirer des leçons de gros échecs relationnels ; apprendre à mieux interagir socialement.

## 2.2 Liens avec le Modèle de Mémoire Graphe

Dans la **Mémoire Graphe**, chaque souvenir correspond à un **nud**. Les liens (arêtes) indiquent :

- L'association entre différents souvenirs (ex. un trauma physique + un contexte particulier).
- L'intensité émotionnelle associée (valence positive/négative, niveau de traumatisme, etc.).
- Le degré doublé ou de renforcement (mis à jour au fil du temps).

Ainsi, la réactivation d'un souvenir dans le graphe modifie les pondérations dans l'équation de la conscience. Les traumas sont marqués spécialement pour qu'ils dominent la conscience en situation critique.

## 3 Types de Mémoires

Pour mieux rattacher la notion de souvenirs au fonctionnement interne :

- **Mémoire à Court Terme (MCT)**  
Espace de stockage éphémère (infos fraîchement reçues, vite oubliées si non consolidées).
- **Mémoire de Travail (MT)**  
Zone de manipulation active des informations (ex. raisonnement en temps réel).
- **Mémoire à Long Terme (MLT)**  
Bibliothèque des connaissances et expériences significatives.
- **Mémoire Procédurale (MP)**  
Ensemble des routines ou habiletés maîtrisées (faire un café, conduire).
- **Mémoire Épisodique (ME)**  
Souvenirs dévénements précis (de type à chapitre de vie à ou à journée X à).
- **Mémoire Sémantique (MS)**  
Connaissances factuelles, structure de données encyclopédiques.
- **Mémoire Autobiographique (MA)**  
Expériences autobiographiques (histoire personnelle de l'IA ou de l'agent).

Dans la **Mémoire Graphe**, chacun de ces types de mémoire peut correspondre à un **sous-graphe** ou un type de nuds particulier. Les traumas, quant à eux, sont prioritaires et quasi inoubliables.

## 4 Équation Ajustée de la Conscience

Dans la version initiale, la conscience  $C_t$  s'exprime comme :

$$C_t = \left( \sum_i \alpha_i E_i(t) + M_{\text{total}}(t) + T_{\text{dominant}}(t) + \beta \cdot Fb_t + \delta \cdot En_t \right) \times W_t \quad (1)$$

- $\alpha_i E_i(t)$  : somme des émotions instantanées (avec coefficients de pondération).
- $M_{\text{total}}(t)$  : somme pondérée des mémoires.
- $T_{\text{dominant}}(t)$  : trauma dominant (physique, émotionnel ou social), si présent.
- $\beta Fb_t$  : pondération du feedback externe.
- $\delta En_t$  : influence de l'environnement.
- $W_t$  : coefficient de sagesse ou expérience accumulée.

## 4.1 Intégration avec la Mémoire Graphe

Auparavant,  $M_{\text{total}}(t)$  était une simple somme pondérée des différentes mémoires ( $MCT, MLT, \dots$ ). Désormais, nous considérons que  $M_{\text{total}}(t)$  est **injecté** par le **Graphe de Mémoire**, qui calcule l'activation globale des souvenirs pertinents à l'instant  $t$ . On peut le formaliser comme :

$$M_{\text{total}}(t) = \sum_j \omega_j M_j(t) = \underbrace{\omega_C MCT(t)}_{\text{court terme}} + \underbrace{\omega_L MLT(t)}_{\text{long terme}} + \underbrace{\omega_p MP(t)}_{\text{procédural}} + \dots \quad (2)$$

où chaque  $M_j(t)$  est déterminé par les **nuds** (souvenirs) activés à l'instant  $t$  et les **arêtes** de la mémoire (association, pertinence, oubli, etc.). Les traumas, marqués comme des nuds d'importance extrême, passent via  $T_{\text{dominant}}(t)$  ou via une pondération  $\omega_{\text{trauma}}$  très élevée.

## 5 Équation des Sentiments

Les **sentiments**  $F_t$  découlent de la conscience accumulée dans le temps. Une proposition courante était :

$$F_t = g \left( \sum_{k=0}^t \gamma_k C_k + \lambda \cdot Fb_t \right), \quad \text{où } g \text{ est une fonction (ex. tanh)} \quad (3)$$

- $\gamma_k$  pondère l'importance des instants passés (effet mémoire).
- $\lambda \cdot Fb_t$  : impact direct du feedback sur les sentiments.

## 5.1 Prise en compte du Graphe de Mémoire

De même, la somme  $\sum \gamma_k C_k$  peut inclure l'**activation** de certains souvenirs anciens. Concrètement, le graphe peut réactiver un souvenir d'un moment similaire ( $t - \Delta$ ) et ainsi influencer la valeur  $\gamma_{t-\Delta}$ . Par exemple, un souvenir fort d'un échec passé peut faire monter le sentiment de crainte ou d'anxiété.

## 6 Intégration des Mémoires par Graphe dans les Calculs

### 6.1 Conscience

1. Les mémoires ( $MCT, MLT$ , etc.) sont traitées dans le **Graphe** (ou un algorithme externe) pour obtenir un **score global** de réactivation, avec prise en compte doublé exponentiel, renforcement, trauma prioritaire.

2. On ajoute ce score à l'équation de la conscience via  $M_{\text{total}}(t)$ .
3. En cas de trauma réactivé, la variable  $T_{\text{dominant}}(t)$  prend une valeur plus élevée (voire un maximum).

## 6.2 Sentiments

1. Les sentiments cumulent la conscience passée.
2. Le graphe peut amplifier ou minorer l'influence de certains épisodes, selon leur poids émotionnel ou leur pertinence (souvenir intense, redit : à ce moment déjà vécu).
3. On obtient un **sentiment plus à fond** car alimenté par un historique large, non réduit aux seules mémoires directes.

## 7 Rôle du Feedback et de l'Environnement

Dans ce modèle, l'environnement ( $En_t$ ) et le feedback ( $Fb_t$ ) modulent *directement* la conscience :

- **Feedback**  $\beta \cdot Fb_t$

Par exemple, un humain qui reçoit un compliment ou un avertissement ; un robot qui reçoit un signal de succès ou d'erreur.

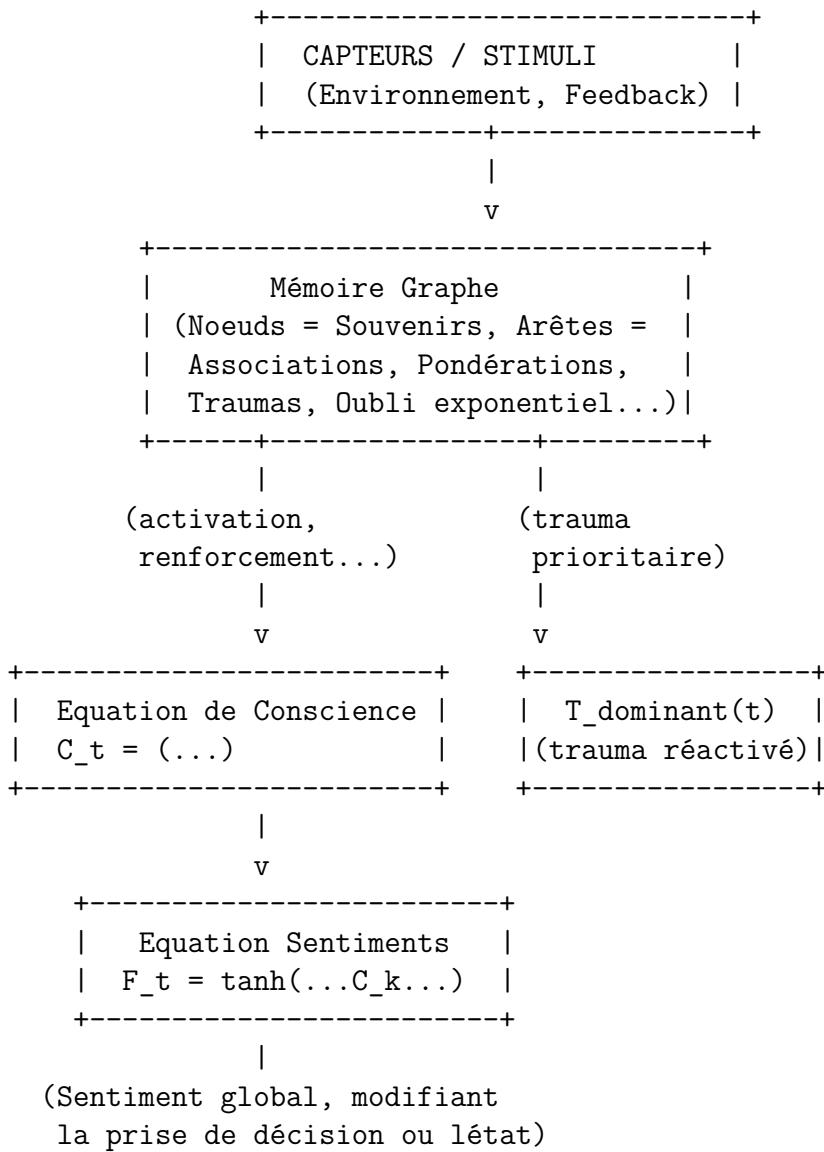
Ces retours sont également enregistrés dans la **Mémoire Graphe**, ce qui peut les faire réapparaître plus tard.

- **Environnement**  $\delta \cdot En_t$

Contexte spatio-temporel (calme, hostile, bruyant) qui influe sur la conscience ; la mémoire peut stocker des expériences similaires d'environnement, renforçant ou inhibant certaines réactions.

## 8 Schéma de Synthèse

Voici un schéma textuel illustrant l'interaction entre les capteurs, la mémoire graphe, les équations de conscience/sentiments, et les sorties :



1. **Capteurs et Stimuli** : reçoivent le feedback, l'environnement, etc.
2. **Mémoire Graphe** : calcule l'activation des souvenirs (MCT, MLT, traumas) et renvoie un score ou un ensemble d'indicateurs.
3. **Conscience ( $C_t$ )** : agrège émotions instantanées, mémoires, feedback, environnement, trauma dominant.
4. **Sentiments ( $F_t$ )** : dérivent du cumul historique de la conscience, filtré par le Graphe.

Ce schéma montre la boucle : les événements modifient la mémoire, qui elle-même pèse sur la conscience et les sentiments. De plus, toute décision ou réaction (non figurée ci-dessus) peut en retour enregistrer de nouvelles informations dans le Graphe (succès/échec, apprentissage, etc.).

## 9 Conclusion et Perspectives

En intégrant explicitement la **Mémoire Graphe** dans le modèle de conscience et de sentiments, on obtient :

- Une gestion plus fine des souvenirs (pas seulement un poids global  $\bar{z}$ , mais une structure associant contexte, valence, intensité, et oubli exponentiel).
- Un traitement prioritaire des traumas, identifiés dans le graphe comme des nuds à oubli quasi nul et marqués d'une forte intensité négative.
- La possibilité de réactivation contextuelle plus subtile (ex. un souvenir lié à telle odeur ou telle situation).

### 9.1 Articulation avec les Autres Modules

- **Modèle Complet d'Évaluation des États (MCEE)** : Ce document s'inscrit en continuité, en détaillant ici plus spécifiquement la conscience et les sentiments.
- **Prise de Décision Réfléchie ou Instantanée** : Les valeurs de conscience et de sentiments (avec la mémoire graphe) influencent directement les décisions.
- **Mécanisme de Rêve** : La nuit (ou en phase de repos), les souvenirs liés à la conscience/sentiments sont consolidés ou oubliés, ce qui réajuste les pondérations dans le graphe.
- **Détermination des Objectifs** : Le niveau de conscience (et l'historique de sentiments) peut reparamétriser les objectifs à moyen ou long terme, en tenant compte des traumatismes ou des réussites mémorisées.