

The background is a dark, textured surface with a complex, glowing pattern. A central, brain-like shape is composed of intricate, swirling lines in shades of blue, purple, and yellow, with numerous bright, star-like points of light. This central shape is surrounded by a network of golden, circuit-like lines that branch out across the frame, resembling a neural network or a complex data structure. The overall effect is one of high-tech, futuristic intelligence.

Introduction au Deep Synergy Learning

Chapitre 1

Mohammed Bouayoun

(bouyao@gmail.com)

Chapitre 1 : Introduction au Deep Synergy Learning (DSL)

| | |
|------------------------------------------------------------------------------------|----|
| Chapitre 1 : Introduction au Deep Synergy Learning (DSL) | 2 |
| 1.1. Contexte et Motivation | 5 |
| 1.1.1. Bref Historique de l'IA | 5 |
| 1.1.2. Émergence des Réseaux Profonds (Deep Learning) et Limites | 6 |
| 1.1.3. Des Approches Conventionnelles à la Synergie Informationnelle | 8 |
| 1.1.4. Positionnement du DSL dans le Panorama de l'IA | 9 |
| 1.1.5. Rôle Potentiel du DSL vers l'IA Forte (IAG)..... | 10 |
| 1.1.6. Objectifs et Contributions Clés du DSL | 12 |
| 1.2. Définitions et Concepts Préliminaires | 14 |
| 1.2.1. Qu'est-ce qu'une "Entité d'Information" ? | 14 |
| 1.2.2. Notion de "Synergie Informationnelle" | 16 |
| 1.2.3. Différence entre Interaction, Synergie et Corrélation | 18 |
| 1.2.4. Approche Hiérarchique vs Approche Auto-Organisée | 21 |
| 1.2.5. Réseaux Neuronaux Traditionnels vs Réseaux Synergiques | 25 |
| 1.2.6. Terminologies Employées dans le DSL | 28 |
| 1.2.7. Exemples Illustratifs de la Synergie dans la Nature | 33 |
| 1.3. Importance de l'Auto-Organisation | 36 |
| 1.3.1. Inspirations Biologiques et Cognitives | 36 |
| 1.3.2. Concepts Clés : Émergence, Auto-Régulation, Feedback | 39 |
| 1.3.3. Comparaison avec les Méthodes d'Apprentissage Classiques | 42 |
| 1.3.4. Le Rôle des Flux d'Information Multimodaux | 45 |
| 1.3.5. Évolution Dynamique et Adaptation en Continu | 48 |
| 1.3.6. Impacts sur la Robustesse et la Résilience des Modèles | 51 |
| 1.3.7. Perspectives pour une Approche plus Globale de l'Apprentissage | 55 |
| 1.4. Architecture Générale du DSL | 59 |
| 1.4.1. Principe de Base : Entités et Liens Synergiques | 59 |
| 1.4.2. Présentation du Synergistic Connection Network (SCN) | 61 |

| | |
|-----------------------------------------------------------------------------|-----|
| 1.4.3. Notion de Cluster et de Macro-Cluster | 63 |
| 1.4.4. Fonctions de Synergie : Distance, Similarité et Co-Information | 67 |
| 1.4.5. Pondérations Adaptatives et Évolution Temporelle | 70 |
| 1.4.6. Interactions Directes et Indirectes | 73 |
| 1.4.7. Synergie binaire et n-aire : au-delà des relations deux à deux | 76 |
| 1.5. Pourquoi une Approche Synergique ? | 79 |
| 1.5.1. Avantages par Rapport aux Réseaux Neuronaux Profonds | 79 |
| 1.5.2. Gestion Naturelle de la Multi-modalité | 84 |
| 1.5.3. Flexibilité vis-à-vis des Données Incomplètes ou Bruitées | 89 |
| 1.5.4. Potentiel d'Auto-Évolution et d'Adaptation Continue | 93 |
| 1.5.5. Réduction de la Dépendance à la Supervision Humaine | 96 |
| 1.5.6. Création de Représentations Riches et plus Interprétables | 99 |
| 1.5.7. Intégration de Dimensions Symboliques ou Cognitives | 104 |
| 1.6. Applications Pressenties et Domaines Impactés | 112 |
| 1.6.1. Vision Artificielle et Reconnaissance d'Objets Complexes | 113 |
| 1.6.2. Analyse Audio et Traitement du Langage Naturel | 116 |
| 1.6.3. Robotique et Systèmes Intelligents Évolutifs | 120 |
| 1.6.4. Recommandation Personnalisée et Systèmes de Décision | 125 |
| 1.6.5. Surveillance, Diagnostic Médical et Anomalies | 129 |
| 1.6.6. Planification et Optimisation dans l'Industrie 4.0 | 133 |
| 1.6.7. Perspectives pour la Recherche Fondamentale en IA Forte | 137 |
| 1.7. Défis, Contraintes et Ouvertures | 141 |
| 1.7.1. Complexité Computationnelle et Scalabilité | 142 |
| 1.7.2. Qualité et Disponibilité des Données | 145 |
| 1.7.3. Développement d'Algorithmes d'Optimisation Appropriés | 148 |
| 1.7.4. Contrôle et Stabilité des Processus Auto-Organisés | 153 |
| 1.7.5. Interprétabilité et Explicabilité pour l'Humain | 157 |
| 1.7.6. Considérations Éthiques et Réglementaires | 160 |
| 1.7.7. Comparaisons Expérimentales avec d'Autres Approches | 164 |

| | |
|--------------------------------------------------------------------------------------|------------|
| 1.8. Positionnement du DSL dans l'Évolution de l'IA | 170 |
| 1.8.1. IA Symbolique vs IA Sub-symbolique : Intégration Potentielle | 171 |
| 1.8.2. DSL et Apprentissage Profond : Collaboration ou Substitution ? | 174 |
| 1.8.3. Approches Hybrides : DSL, RL (Reinforcement Learning) et Logique | 177 |
| 1.8.4. Rôle de la Mémoire et de l'Attention dans le DSL | 181 |
| 1.8.5. Tendances Futures : Vers une IA Forte, Consciente ? | 185 |
| 1.8.6. Effet Sur la Recherche Interdisciplinaire | 189 |
| 1.8.7. Exemple de Convergence : DSL & Neurosciences | 192 |

1.1. Contexte et Motivation

1.1.1. Bref Historique de l'IA

L'intelligence artificielle (IA) constitue un domaine de recherche qui s'inscrit dans une longue tradition, remontant à la première moitié du XXe siècle. Son évolution a été marquée par des jalons fondamentaux qui ont façonné la discipline telle qu'elle se présente aujourd'hui. Pour mieux appréhender l'émergence de l'Apprentissage Profond Synergétique (APS), ou **Deep Synergy Learning (DSL)** en anglais, il est essentiel d'examiner les principaux courants qui ont successivement vu le jour ou cohabité.

Des précurseurs conceptuels à la naissance de l'IA (années 1940–1950)

Les premières réflexions autour de l'intelligence artificielle trouvent leur origine dans les travaux d'Alan Turing, notamment à travers son article de 1950, *Computing Machinery and Intelligence*. Dans ce texte fondateur, il introduit le célèbre test de Turing, conçu pour évaluer la capacité d'une machine à simuler l'intelligence humaine dans le cadre d'une interaction textuelle.

C'est lors de la conférence de Dartmouth en 1956 que l'IA est officiellement définie comme un champ scientifique. Cette rencontre, organisée par John McCarthy, Marvin Minsky, Claude Shannon et Nathan Rochester, marque l'acte fondateur de la discipline. Elle établit les bases conceptuelles et méthodologiques nécessaires à la création de systèmes intelligents.

L'âge d'or de l'IA symbolique (années 1950–1970)

Les premières approches de l'IA reposaient sur la manipulation de symboles et de règles logiques pour modéliser la pensée humaine. Ces travaux, regroupés sous l'appellation **IA symbolique**, cherchaient à reproduire des processus cognitifs à travers des systèmes de règles formelles. Des programmes tels que *Logic Theorist* (1956) et *General Problem Solver* (1959) illustrent cette démarche en appliquant des algorithmes de recherche pour résoudre des problèmes mathématiques ou généraux.

Cependant, bien que performante dans des environnements fermés, l'IA symbolique montrait des limites lorsqu'il s'agissait de traiter des situations complexes ou incertaines. Cette incapacité à généraliser au-delà de scénarios spécifiques a révélé la nécessité d'approches complémentaires.

L'essor puis l'hiver des réseaux de neurones (années 1960–1980)

Parallèlement à l'IA symbolique, le courant connexionniste, basé sur les réseaux de neurones artificiels, gagne en popularité. Des chercheurs comme Frank Rosenblatt, Bernard Widrow et Marcian Hoff ont exploré le potentiel des réseaux d'unités élémentaires capables d'apprendre à partir de données, comme en témoigne le Perceptron (1958). Cependant, les travaux de Marvin Minsky et Seymour Papert, publiés en 1969, ont mis en lumière les limites des réseaux de neurones monocouches, incapables de traiter des problèmes non linéaires. Cette critique, combinée à l'absence de méthodes efficaces pour entraîner des réseaux multicouches, a conduit au déclin temporaire de l'intérêt pour le connexionnisme, marquant le premier "hiver de l'IA".

L'approche statistique et le renouveau du connexionnisme (années 1980–1990)

Le milieu des années 1980 voit la renaissance des réseaux de neurones grâce à la découverte de la rétropropagation, une méthode permettant d'entraîner efficacement des réseaux multicouches. Ce renouveau est accompagné d'une montée en puissance des approches statistiques de l'apprentissage automatique, telles que les modèles graphiques probabilistes et les machines à vecteurs de support (SVM). Ces outils élargissent les applications de l'IA à des domaines variés, notamment la vision par ordinateur et le traitement du langage naturel.

L'explosion du Deep Learning (années 2000–2010)

La fin des années 2000 marque l'avènement du Deep Learning, rendu possible par la convergence de trois facteurs majeurs. L'abondance des données massives issues d'Internet et des technologies numériques a offert un terrain fertile pour l'entraînement des modèles. L'émergence de ressources de calcul performantes, telles que les GPU et les TPU, a permis de traiter des volumes de données colossaux. Enfin, des progrès méthodologiques, incluant de nouvelles fonctions d'activation et des stratégies d'optimisation, ont renforcé l'efficacité des réseaux de neurones profonds. Ces avancées ont conduit à des percées significatives dans des domaines comme la reconnaissance d'images, la traduction automatique et les jeux stratégiques, illustrées notamment par le succès d'AlphaGo.

Malgré ces réussites, le Deep Learning présente certaines limites, notamment une forte dépendance aux données annotées, un manque d'interprétabilité et des difficultés à généraliser. Ces contraintes mettent en évidence la nécessité d'approches complémentaires.

Vers la synergie informationnelle et l'IA forte

Face aux limitations des méthodes actuelles, de nouveaux paradigmes ont émergé, mettant l'accent sur l'auto-organisation et les interactions dynamiques entre sources d'information. Le **Deep Synergy Learning (DSL)** s'inscrit dans cette dynamique en intégrant des notions de coopération et de coévolution au sein des systèmes intelligents. Cette approche vise à dépasser les modèles existants en favorisant une intelligence plus adaptable et autonome, ouvrant ainsi la voie à l'intelligence artificielle générale (IAG).

1.1.2. Émergence des Réseaux Profonds (Deep Learning) et Limites

L'émergence des réseaux profonds, ou **Deep Learning**, constitue un tournant fondamental dans l'histoire de l'intelligence artificielle. Bien que les réseaux de neurones artificiels aient été introduits dès les années 1960, ce n'est qu'au début des années 2010 qu'ils se sont imposés comme une technologie centrale grâce à plusieurs avancées majeures. Ces progrès incluent l'accroissement massif des données numériques, le développement de ressources matérielles puissantes comme les GPU, et l'introduction de méthodologies innovantes. Cependant, malgré ces succès, le Deep Learning continue de présenter des limitations importantes qui suscitent l'exploration de nouvelles approches.

L'essor du Deep Learning : facteurs clés

L'avènement du Deep Learning repose sur plusieurs éléments déterminants. D'une part, l'ère numérique a produit une quantité sans précédent de données issues de diverses sources, telles que

des images, des vidéos, des textes ou des signaux. Ces ensembles de données volumineux, comme le dataset ImageNet, ont permis d'entraîner des réseaux plus profonds et plus complexes. D'autre part, l'utilisation des GPU, conçus initialement pour le traitement graphique, a permis d'accélérer les calculs massivement parallèles nécessaires à l'entraînement des modèles. Ces avancées matérielles ont été accompagnées de progrès méthodologiques, comme l'introduction de fonctions d'activation avancées, la régularisation par dropout ou encore la normalisation de lots (batch normalization). Enfin, les succès visibles dans des applications telles que la reconnaissance d'images, la traduction neuronale et les jeux stratégiques ont renforcé l'intérêt pour cette technologie.

Les forces des réseaux profonds

Les réseaux de neurones profonds ont apporté des avantages significatifs dans le traitement des données complexes. L'une de leurs principales forces réside dans leur capacité à extraire automatiquement des caractéristiques pertinentes à partir des données brutes, ce qui simplifie considérablement les tâches d'ingénierie des features. De plus, leur performance dans des domaines tels que la classification d'images ou la reconnaissance vocale dépasse de loin celle des méthodes traditionnelles. Lorsqu'ils sont entraînés avec suffisamment de données pertinentes, ces réseaux montrent une certaine aptitude à généraliser en identifiant des motifs complexes que d'autres approches ne peuvent pas détecter.

Limites et défis persistants

Malgré leurs performances impressionnantes, les réseaux profonds présentent plusieurs limites. Tout d'abord, leur dépendance aux données annotées est problématique, car l'annotation de grands volumes de données est coûteuse et parfois impraticable dans certains contextes. Ensuite, leur nature de boîte noire rend difficile l'interprétation des décisions prises, ce qui pose des problèmes d'explicabilité, notamment dans des secteurs critiques comme la santé ou la justice. Les réseaux profonds sont également vulnérables aux attaques adversariales, où des perturbations mineures peuvent induire des erreurs majeures. De plus, leur capacité de généralisation reste limitée lorsqu'ils sont confrontés à des contextes différents de ceux rencontrés durant l'apprentissage. Enfin, le coût énergétique élevé de l'entraînement de ces modèles soulève des préoccupations en termes de durabilité et d'accessibilité.

Perspectives et nouveaux paradigmes

Face à ces limitations, de nouvelles approches visent à repousser les frontières du Deep Learning. L'apprentissage peu supervisé ou auto-supervisé, qui réduit la nécessité de données annotées, est en plein essor. Par ailleurs, l'intégration du raisonnement symbolique avec les réseaux de neurones, dans une approche neuro-symbolique, pourrait permettre une intelligence plus transparente et explicable. Enfin, des recherches explorent des architectures auto-organisées où les informations interagissent de manière dynamique, rompant avec la rigidité des modèles actuels. Ces évolutions pourraient ouvrir la voie à des systèmes plus robustes, adaptatifs et proches d'une véritable intelligence artificielle générale.

1.1.3. Des Approches Conventionnelles à la Synergie Informationnelle

L'histoire de l'intelligence artificielle témoigne de la succession de grandes tendances méthodologiques, chacune cherchant à reproduire ou modéliser certains aspects de l'intelligence humaine ou animale. Parmi ces approches, les approches symboliques et les approches statistiques ou connexionnistes ont dominé. Les premières s'appuient sur la formalisation explicite de la connaissance à travers des règles et des faits, tandis que les secondes reposent sur l'apprentissage à partir de grandes quantités de données via des modèles tels que les réseaux de neurones artificiels. Bien que ces paradigmes aient permis des avancées considérables, ils montrent leurs limites lorsqu'il s'agit de traiter des systèmes complexes, dynamiques ou émergents. La synergie informationnelle s'inscrit dans une tentative de dépasser ces restrictions en proposant une approche inspirée des systèmes naturels.

Inspiration des systèmes complexes et de la nature

Les systèmes naturels, tels que les écosystèmes, les colonies d'insectes ou le cerveau humain, illustrent un principe fondamental selon lequel l'ensemble dépasse la somme de ses parties. Ces systèmes se caractérisent par l'auto-organisation, où des structures émergent sans planification globale, la robustesse grâce à la redondance des composants, et la capacité d'adaptation en réponse à des variations environnementales. Les entités locales interagissent dynamiquement à travers des signaux ou des mécanismes simples, conduisant à des comportements collectifs complexes. Transposé à l'intelligence artificielle, ce principe propose de considérer les informations comme des entités actives capables d'interagir, de coopérer et de s'adapter en permanence.

Limites des approches conventionnelles

Les approches classiques du Deep Learning, bien qu'efficaces, restent enfermées dans une structure hiérarchique rigide où les données circulent de manière linéaire ou avec des boucles limitées. Cela restreint leur capacité à modéliser des interactions complexes ou des phénomènes émergents. En particulier, dans des scénarios multi-modaux combinant des flux hétérogènes comme la vision, le texte ou l'audio, l'intégration repose souvent sur des mécanismes de fusion statiques qui ne permettent pas une réelle interaction entre les modalités. Ces architectures peinent également à s'adapter à des environnements dynamiques ou partiellement inconnus.

Concept de synergie informationnelle

La synergie informationnelle propose une vision radicalement différente. Les entités d'information, au lieu d'être des entrées passives dans un modèle prédéfini, deviennent des acteurs autonomes capables d'interagir et d'évoluer. Elles peuvent évaluer leur synergie avec d'autres entités, c'est-à-dire mesurer leur capacité à produire ensemble des informations plus riches que la somme de leurs contributions individuelles. En fonction de cette évaluation, elles peuvent nouer, rompre ou modifier des connexions, favorisant l'émergence de structures auto-organisées et dynamiques. Ce processus dépasse la simple adaptation des pondérations traditionnelles en permettant des changements topologiques dans les réseaux, où de nouvelles entités peuvent apparaître, fusionner ou disparaître.

Vers le Deep Synergy Learning (DSL)

Le **Deep Synergy Learning (DSL)** s'inscrit dans cette continuité en cherchant à unifier et enrichir les paradigmes existants. Alors que le Deep Learning met l'accent sur la profondeur des

représentations, le DSL valorise la richesse et la dynamique des interactions entre les entités d'information. Les bénéfices attendus incluent une meilleure intégration des modalités multiples, une résilience accrue face à l'incertitude et une capacité d'apprentissage continu où le réseau peut évoluer en réponse à de nouvelles données sans nécessiter un réapprentissage complet. En favorisant des interactions flexibles et adaptatives, le DSL ambitionne de jeter les bases d'une intelligence artificielle plus proche des capacités cognitives humaines.

1.1.4. Positionnement du DSL dans le Panorama de l'IA

Le **Deep Synergy Learning (DSL)** s'inscrit dans la continuité des avancées en intelligence artificielle, tout en introduisant une approche novatrice fondée sur la coopération dynamique et l'auto-organisation des entités d'information. Pour mieux comprendre son originalité, il convient de le situer par rapport aux grandes tendances historiques et actuelles, à savoir l'**IA symbolique**, le **connexionnisme** incarné par le **Deep Learning**, et les **approches hybrides ou émergentes**.

Héritage des approches symboliques et connexionnistes

Le DSL se distingue des systèmes classiques d'IA symbolique, qui reposent sur la formalisation explicite de règles logiques et de bases de connaissances, en permettant l'émergence spontanée de structures et d'organisations adaptées au contexte. Contrairement à la rigidité des approches symboliques, il s'appuie sur des mécanismes auto-adaptatifs qui modifient les interactions en fonction des données disponibles. Par ailleurs, il prolonge l'héritage connexionniste en mettant l'apprentissage au centre de son fonctionnement. Alors que le Deep Learning repose sur des architectures statiques et hiérarchiques comme les CNN ou Transformers, le DSL introduit une dynamique d'interaction entre les entités, permettant une réorganisation continue des connexions en fonction de leur synergie.

Inspiration des systèmes complexes et émergents

Le DSL puise son inspiration dans les systèmes naturels complexes tels que les écosystèmes ou les réseaux neuronaux biologiques. Ces systèmes se caractérisent par des propriétés globales qui émergent de l'interaction locale d'éléments simples, sans planification centrale. Dans cette perspective, chaque entité d'information dans le DSL peut établir, rompre ou renforcer des connexions en fonction de la synergie qu'elle partage avec d'autres entités. Cette dynamique dépasse les schémas hiérarchiques classiques en introduisant des boucles de rétroaction et en intégrant des structures internes évolutives. Contrairement aux approches traditionnelles où les flux d'information sont souvent indépendants et combinés à un stade tardif, le DSL favorise un tissage constant des flux, améliorant ainsi la robustesse et l'adaptabilité.

Positionnement dans l'écosystème des approches hybrides

Le DSL se distingue également des approches neuro-symboliques qui intègrent explicitement des règles logiques dans des systèmes neuronaux. Bien qu'il ne repose pas directement sur des formalismes logiques, il permet l'émergence de règles implicites via des interactions coopératives et des clusters auto-organisés. De plus, il participe à la quête d'un apprentissage continu en permettant l'évolution permanente des connexions et en favorisant une réorganisation constante en réponse aux nouveaux contextes. Cela en fait une solution prometteuse pour surmonter le problème de l'oubli catastrophique souvent rencontré dans l'apprentissage automatique.

Contribution unique à l'intelligence générale

L'IA forte, ou intelligence générale, se caractérise par sa capacité à s'adapter de manière flexible à des environnements variés et non balisés. Le DSL offre une contribution unique à cet objectif en introduisant le principe de coévolution des entités d'information. Cette adaptabilité intrinsèque repose sur l'auto-organisation et la synergie, qui sont considérées comme des éléments essentiels pour l'émergence d'une intelligence créative et autonome. Plutôt que de remplacer le Deep Learning, le DSL peut s'y intégrer en tant que niveau supplémentaire, enrichissant les représentations neuronales classiques par une interactivité et une plasticité accrues.

1.1.5. Rôle Potentiel du DSL vers l'IA Forte (IAG)

L'**Intelligence Artificielle Forte (IAG)**, également désignée par intelligence artificielle générale, se définit comme la capacité d'une machine à réaliser des tâches cognitives humaines de manière autonome, flexible et adaptative, en s'appuyant sur une compréhension et un apprentissage comparables à ceux d'un être humain. Cet objectif ultime de l'intelligence artificielle reste encore largement théorique, malgré les avancées impressionnantes des systèmes actuels, notamment ceux basés sur le **Deep Learning**. Ces modèles, bien que performants dans des tâches spécialisées, peinent à atteindre le niveau de polyvalence et de plasticité requis pour prétendre à une véritable IA générale. Le **Deep Synergy Learning (DSL)** se positionne comme une approche novatrice et prometteuse, offrant un cadre conceptuel pour dépasser les limites des modèles traditionnels.

Repenser la Généralité en IA

Une intelligence générale ne se limite pas à exceller dans une tâche unique, mais doit pouvoir évoluer dans des environnements variés et imprévisibles. Le DSL se distingue par son aptitude à favoriser l'apprentissage multi-contextuel grâce à ses mécanismes d'**auto-organisation** et de **synergie informationnelle**. Ces mécanismes permettent au système de s'adapter en temps réel à des situations inédites, en reconfigurant les relations entre entités d'information. Contrairement aux systèmes conventionnels, qui restent figés dans leurs architectures ou nécessitent une ingénierie lourde pour évoluer, le DSL peut ajuster sa structure de manière dynamique.

La richesse des interactions entre les entités est une autre caractéristique essentielle du DSL. Plutôt que de juxtaposer des modules spécialisés sans réelle interaction, il propose une approche où chaque flux d'information, qu'il soit visuel, textuel ou sensoriel, interagit et co-évolue avec les autres. Ce type d'interconnexion fluide permet l'émergence de représentations conceptuelles globales, essentielles pour traiter la complexité du monde réel.

Enfin, le DSL intègre une dimension d'**adaptabilité évolutive**, permettant au système de non seulement apprendre, mais également désapprendre, réviser et transformer ses structures en fonction de nouvelles expériences. Cette adaptabilité dépasse celle des architectures classiques, qui, bien qu'elles ajustent leurs pondérations, conservent généralement une topologie statique.

Auto-Organisation et Plasticité Dynamique

L'un des principes fondamentaux du DSL est l'**auto-organisation**, inspirée des systèmes biologiques. À l'image du cerveau humain, il privilégie une architecture ouverte, où les entités d'information ne suivent pas une trajectoire linéaire prédéfinie, mais interagissent selon des

schémas dynamiques et adaptatifs. Cette capacité à reconfigurer les flux internes permet au système de s'auto-réguler et d'évoluer en fonction des besoins contextuels.

La **plasticité dynamique** du DSL constitue un autre levier crucial pour atteindre l'IA forte. Elle dépasse l'ajustement traditionnel des pondérations en introduisant la possibilité de créer de nouvelles connexions, d'en supprimer d'obsolètes ou de réorganiser des clusters entiers. Cette plasticité ouvre la voie à une exploration continue et à une innovation constante, des caractéristiques indispensables pour développer une intelligence véritablement générale.

De la Spécialisation à la Flexibilité Généralisée

Les systèmes d'apprentissage profond ont démontré leur capacité à exceller dans des domaines spécifiques, comme la reconnaissance d'images ou la traduction automatique. Cependant, cette spécialisation s'accompagne d'une incapacité à transférer efficacement les connaissances d'un domaine à un autre. Le DSL propose une alternative en favorisant une **coopération dynamique** entre différentes modalités ou tâches, permettant une synergie globale entre sous-systèmes.

Ce paradigme ouvre la possibilité d'un transfert de connaissances fluide, où les acquis d'une modalité, comme l'analyse visuelle, enrichissent l'apprentissage dans une autre, telle que la compréhension du langage. Cette capacité de réutilisation et de réorganisation constitue une avancée majeure vers une **flexibilité généralisée**, essentielle pour le développement de l'IAG.

Vers une Architecture Cognitive Unifiée

Le cerveau humain se distingue par son haut degré d'interconnexion et sa capacité à faire émerger des significations à partir de l'interaction entre différentes aires sensorielles et associatives. Le DSL s'inspire de ce modèle en proposant une architecture où les interactions dynamiques entre les entités d'information jouent un rôle central. Cette approche favorise non seulement l'émergence de comportements complexes, mais également celle de **schémas cognitifs** riches et adaptatifs, proches de ceux observés dans les processus humains.

De plus, le DSL vise à dépasser la simple corrélation statistique en permettant l'émergence de concepts et de significations. Les interactions entre entités d'information, structurées par des mécanismes d'auto-organisation, contribuent à la construction de **nœuds sémantiques**, amorçant ainsi une véritable compréhension du monde environnant.

Défis et Perspectives

Bien que le DSL représente une avancée prometteuse, plusieurs défis doivent être relevés pour qu'il puisse concrétiser son potentiel en tant que moteur de l'IAG. L'**efficacité algorithmique** constitue un premier enjeu, car les processus d'auto-organisation et de réorganisation dynamique peuvent être coûteux en termes de calcul. Trouver des solutions d'optimisation est donc essentiel.

L'**équilibre entre stabilité et plasticité** est également un aspect critique. Le système doit éviter de se réorganiser de manière excessive, ce qui pourrait nuire à la cohérence de ses apprentissages, tout en maintenant une capacité suffisante d'adaptation.

Enfin, les questions d'**interprétabilité** et de **sécurité** demeurent cruciales. Une architecture aussi dynamique et complexe pourrait poser des problèmes de traçabilité et de contrôle, notamment dans des domaines sensibles comme la santé ou la finance.

1.1.6. Objectifs et Contributions Clés du DSL

Le **Deep Synergy Learning (DSL)** s'inscrit comme une évolution fondamentale dans le domaine de l'**intelligence artificielle (IA)**, mettant en avant une **approche dynamique** et hautement **adaptative** pour surmonter les limites des architectures hiérarchiques classiques, telles que celles du **Deep Learning**. Cette nouvelle méthodologie repose sur l'idée centrale de **synergie informationnelle**, où les entités d'information interagissent et s'organisent de manière autonome pour produire des représentations plus riches et efficaces. L'objectif principal du DSL est d'intégrer les principes issus du **connexionnisme** et des **systèmes complexes**, tout en introduisant des mécanismes avancés d'**auto-organisation** et d'**apprentissage évolutif**.

A. Cadre théorique et objectifs fondamentaux

Le DSL repose sur une modélisation qui privilégie les interactions **synergiques** entre les entités d'information. Contrairement aux architectures fixes, où les flux d'information sont généralement dirigés et figés, le DSL permet une **reconfiguration continue** des liens internes. Cette flexibilité est rendue possible grâce à des mécanismes d'**évaluation dynamique** de la synergie, notée $S(i, j)$, entre deux entités i et j . Mathématiquement, la synergie entre deux entités peut être décrite comme une fonction dépendant de leurs représentations respectives x_i et x_j :

$$S(i, j) = \Phi(x_i, x_j),$$

où Φ est une fonction d'évaluation qui peut être basée sur la corrélation, l'entropie conjointe, ou toute autre mesure adaptée au contexte.

L'objectif fondamental est de maximiser cette synergie globale dans le système, définie comme une somme pondérée des synergies individuelles :

$$\text{Synergie totale} = \sum_{i \neq j} w_{ij} S(i, j),$$

où w_{ij} représente les poids ajustables des connexions entre i et j . En optimisant cette fonction, le DSL favorise des **interactions pertinentes** et élimine les connexions inutiles ou redondantes, améliorant ainsi l'efficacité computationnelle et la robustesse des représentations générées.

L'un des objectifs clés du DSL est également d'offrir une plateforme pour l'**intégration multimodale**. En incorporant des données issues de différentes modalités, comme le texte, les images ou les signaux temporels, le système est capable d'unifier ces informations de manière **contextuelle**. Chaque modalité contribue activement à l'apprentissage global, tout en bénéficiant des enrichissements apportés par les autres, dans un cadre coopératif.

B. Contributions principales et innovations méthodologiques

Le DSL propose une contribution notable en introduisant un mécanisme d'**auto-organisation** et de **plasticité structurelle**. Contrairement aux approches classiques où les architectures sont déterminées à l'avance, le DSL permet une **reconfiguration dynamique** des connexions entre entités, suivant des critères d'efficacité synergiques. Cette adaptabilité repose sur un processus

d'optimisation continue où chaque entité ajuste ses liens en fonction de la pertinence des informations échangées. Ce mécanisme peut être formalisé comme suit :

$$\frac{\partial w_{ij}}{\partial t} = \alpha \cdot \nabla S(i, j),$$

où α est un facteur d'apprentissage, et $\nabla S(i, j)$ désigne le gradient de la synergie entre i et j .

Une autre innovation majeure réside dans la gestion des **clusters auto-organisés**, qui sont des regroupements spontanés d'entités hautement synergiques. Ces clusters peuvent émerger, se fusionner ou se dissoudre en fonction des interactions observées, offrant une **modularité naturelle** au système. Cette capacité de création et de dissolution des clusters renforce la **résilience** et l'**évolutivité**, deux qualités cruciales pour les environnements complexes et dynamiques.

Enfin, le DSL se distingue par sa capacité à capturer des **représentations sémantiques profondes**, en favorisant l'émergence de **patrons globaux** à partir d'interactions locales. Cette propriété rapproche le DSL des processus cognitifs humains, où la compréhension découle souvent d'une synthèse d'informations contextuelles et émergentes.

C. Avantages du DSL

Le DSL présente des avantages significatifs par rapport aux paradigmes d'apprentissage traditionnels. L'un des principaux bénéfices est sa capacité à **maximiser la synergie informationnelle**, ce qui améliore la richesse et la robustesse des représentations apprises. Cette optimisation synergique permet au système d'exploiter pleinement la diversité des données, y compris dans des environnements multi-modaux. Par ailleurs, la **plasticité dynamique** du DSL garantit une adaptabilité exceptionnelle face à des scénarios évolutifs ou imprévisibles, tout en réduisant les besoins en supervision humaine. En outre, son cadre auto-organisé favorise une meilleure **résilience** aux perturbations, grâce à la redondance et à la flexibilité des connexions internes.

D. Limites et défis à surmonter

Malgré ses nombreux atouts, le DSL n'est pas exempt de défis. L'un des obstacles majeurs est la **complexité computationnelle** inhérente à l'évaluation continue des synergies et à la reconfiguration des connexions. Cette complexité pourrait limiter son application à grande échelle, en particulier dans des contextes où les ressources en calcul sont contraintes. De plus, la nature hautement dynamique et adaptative du DSL pose des **problèmes d'interprétabilité**, car il devient difficile de tracer ou d'expliquer précisément les décisions prises par le système. Enfin, le maintien d'un **équilibre entre stabilité et plasticité** est une question délicate. Une plasticité excessive peut provoquer une instabilité, tandis qu'une structure trop rigide risque de nuire à l'efficacité globale.

1.2. Définitions et Concepts Préliminaires

Dans cette partie, nous allons introduire les notions fondamentales qui sous-tendent le **Deep Synergy Learning (DSL)**. Alors que la section précédente (1.1) portait sur le contexte historique, les motivations et la place du DSL dans l'écosystème de l'IA, le présent segment (1.2) est consacré à la **structure conceptuelle** du DSL et aux **fondements** qui permettent d'en comprendre les mécanismes profonds.

Nous débuterons par la définition de ce qu'est une **"Entité d'Information"** (1.2.1), pierre angulaire du paradigme DSL. Nous poursuivrons en expliquant le concept de **"Synergie Informationnelle"** (1.2.2) et en clarifiant la différence entre interaction, synergie et simple corrélation (1.2.3). Nous verrons ensuite comment le DSL se démarque d'une **approche strictement hiérarchique** (1.2.4) et en quoi il diverge des **réseaux neuronaux traditionnels** (1.2.5). Enfin, nous préciserons la **terminologie** spécifique (1.2.6) et illustrerons les principes de synergie par des **exemples concrets** (1.2.7).

1.2.1. Qu'est-ce qu'une "Entité d'Information" ?

Le concept d'**entité d'information** est l'un des piliers fondamentaux du **Deep Synergy Learning**. Dans le **DSL**, une entité d'information ne se réduit pas à un simple point de données isolé, mais représente un véritable **objet d'apprentissage** doté de plusieurs propriétés.

Elle possède d'abord une **représentation interne**, qui peut prendre la forme d'un vecteur, d'un tenseur ou d'une fonction. Elle présente également des **caractéristiques dynamiques**, lui permettant d'évoluer, de se modifier et d'entretenir des liens synergiques avec d'autres entités. Enfin, elle dispose d'un **historique** ou d'une mémoire partielle de ses interactions passées, influençant ainsi son comportement futur et sa capacité d'adaptation.

Pour formaliser une entité d'information \mathcal{E} , on se place dans un espace vectoriel (ou parfois un espace de Hilbert plus général) :

$$\mathcal{E}_k \in \mathbb{R}^d, \quad \text{ou} \quad \mathcal{E}_k \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_p},$$

suivant la nature des données (vecteur, matrice, tenseur, etc.). Par exemple, une image peut être encodée en tant que tenseur 3D (hauteur \times largeur \times canaux de couleur), tandis qu'un signal audio pourra être représenté sous forme de séries temporelles dans \mathbb{R}^d .

Dans certains cas, la représentation peut également être probabiliste. Ainsi, \mathcal{E}_k peut être décrite par une distribution de probabilité $\mathcal{P}_k(\mathbf{x})$ sur un certain espace $\mathbf{x} \in \mathcal{X}$. L'important est de conserver la possibilité de **mesurer** la distance, la similarité ou la divergence entre deux entités :

$$\text{dist}(\mathcal{E}_k, \mathcal{E}_m) \quad \text{ou} \quad \text{sim}(\mathcal{E}_k, \mathcal{E}_m).$$

Outre la représentation brute, une entité peut avoir des **paramètres internes** (poids, biais, etc.) qui se modifient selon le temps ou selon les interactions :

$$\theta_k = \{\theta_{k,1}, \theta_{k,2}, \dots, \theta_{k,\ell}\}.$$

Ces paramètres influent sur le “comportement” de l’entité, c’est-à-dire sa manière de calculer des **scores de similarité** ou des **fonctions de sortie**. On peut aussi décrire un **état** interne $\mathbf{s}_k(t)$ évoluant avec t , le temps (ou la phase d’apprentissage) :

$$\mathbf{s}_k(t) \in \mathbb{R}^d,$$

indiquant, par exemple, le niveau de confiance ou les caractéristiques discriminantes apprises jusqu’à l’instant t . Cet état peut servir de base à la mise à jour des **connexions synergiques** entre l’entité \mathcal{E}_k et d’autres entités \mathcal{E}_m .

Dans le cadre du DSL, nous pouvons associer à chaque entité \mathcal{E}_k un **ensemble** de composants :

$$\mathcal{E}_k = (\mathbf{x}_k, \mathbf{s}_k, \theta_k, \dots),$$

où :

- $\mathbf{x}_k \in \mathbb{R}^d$ est la représentation courante (ex. : vecteur de caractéristiques, image encodée).
- \mathbf{s}_k est l’état interne dynamique (optionnel ou modulable).
- θ_k représente des **paramètres d’ajustement** ou d’apprentissage.
- D’autres composants pourraient inclure la **mémoire** (historique), les **métadonnées**, etc.

Cette formulation a pour but de **généraliser** la notion de “neurone” ou de “vecteur de données” pour en faire une entité d’apprentissage **active** et **adaptative**, au cœur des mécanismes d’interaction synergiques.

Pour illustrer concrètement la notion d’entité d’information, considérons la tâche de **reconnaissance de situations** dans une vidéo associée à un flux audio. Dans ce scénario, l’entité $\mathcal{E}_{\text{visuelle}}$ peut représenter un **descripteur d’image** extrait par un réseau de neurones convolutionnel (CNN), enrichi de **paramètres** relatifs à la forme ou à la pose des objets détectés. De son côté, l’entité $\mathcal{E}_{\text{auditive}}$ peut regrouper une **carte d’intensité fréquentielle** (spectrogramme) et un **état** décrivant la tonalité ou le niveau de bruit ambiant.

Ces deux entités ne sont pas de simples **blocs de données isolés**. Elles sont conçues pour **interagir**, **se synchroniser** ou même **fusionner** lorsque leur **synergie** (au sens de la section 1.2.2) atteint un niveau suffisamment élevé. Autrement dit, si les informations issues des canaux visuel et auditif s’enrichissent mutuellement, elles ont la possibilité de renforcer leurs liens et, potentiellement, de s’intégrer au point de former une entité commune, apte à traiter des signaux audiovisuels de manière coordonnée. Cette démarche souligne la **plasticité** du Deep Synergy Learning, qui autorise une reconfiguration permanente des relations entre entités pour améliorer la représentation globale de la scène.

En définitive, l’entité d’information constitue le **nœud élémentaire** du DSL. C’est à **travers** elle et **par** elle que les mécanismes de synergie prennent forme, permettant l’émergence de structures d’apprentissage plus complexes. Le **design** même de chaque entité, qu’il s’agisse de sa représentation (vecteur, tenseur, distribution de probabilité), de son état ou de ses paramètres d’ajustement, détermine directement l’**expressivité** et l’**efficacité** de l’apprentissage au sein du réseau. Le choix judicieux de ces attributs, adapté à la modalité (vision, audio, texte, etc.) et à la

tâche visée, facilite la formation de **liaisons synergiques** fructueuses et, par conséquent, contribue à la robustesse et à la performance globale du système DSL.

1.2.2. Notion de “Synergie Informationnelle”

L’un des concepts centraux du Deep Synergy Learning (DSL) est la **synergie informationnelle**. Il s’agit de la capacité de deux (ou plusieurs) entités d’information à produire, ensemble, un **contenu** ou une **performance** impossible à atteindre (ou significativement moins bonne) si elles agissaient de manière isolée. Dans un cadre mathématique, la synergie se formalise par une mesure qui évalue l’**apport mutuel** entre les entités. Plus cette mesure est élevée, plus les entités concernées s’enrichissent mutuellement, amplifiant leur pouvoir de représentation ou de décision.

Définition générale. Considérons deux entités d’information, \mathcal{E}_i et \mathcal{E}_j . En première approximation, on peut définir la **synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$ comme une fonction qui quantifie à quel point la prise en compte conjointe de \mathcal{E}_i et de \mathcal{E}_j **améliore** un critère d’apprentissage (la prédiction d’une variable cible, la qualité d’une représentation, etc.).

$$S(\mathcal{E}_i, \mathcal{E}_j) = f(\mathcal{E}_i, \mathcal{E}_j),$$

où $f(\cdot)$ peut être :

- Une **mesure d’entropie conjointe** (ou de co-information) en théorie de l’information,
- Un **gain de performance** par rapport à une référence (p. ex. différence de log-vraisemblance),
- Une **fonction de similarité/distance** qui prend en compte des aspects non linéaires et adaptatifs.

Cette fonction f doit être conçue pour refléter la notion que “le tout est plus que la somme des parties”. Ainsi, il est d’usage de considérer qu’une **haute synergie** indique que l’association $\{\mathcal{E}_i, \mathcal{E}_j\}$ est nettement plus informative que chacune des entités prise isolément.

En théorie de l’information, on peut s’appuyer sur l’**entropie conjointe** et la **co-information**. Par exemple, si \mathbf{X}_i et \mathbf{X}_j sont les variables aléatoires (représentant respectivement les entités \mathcal{E}_i et \mathcal{E}_j), on définit :

$$I(\mathbf{X}_i; \mathbf{X}_j) = H(\mathbf{X}_i) + H(\mathbf{X}_j) - H(\mathbf{X}_i, \mathbf{X}_j),$$

où $H(\cdot)$ est l’entropie (de Shannon, ou d’autres formes d’entropie plus générales). Cette quantité $I(\mathbf{X}_i; \mathbf{X}_j)$ mesure l’**information mutuelle** entre \mathbf{X}_i et \mathbf{X}_j . Toutefois, l’information mutuelle standard ne distingue pas toujours la **synergie** de la **redondance**.

Pour caractériser la synergie stricto sensu, plusieurs travaux de théorie de l’information proposent des mesures de **co-information** plus élaborées, voire des “Partial Information Decomposition” (PID), qui visent à séparer la part de redondance et la part de synergie :

$$\text{Synergie}(\mathbf{X}_i, \mathbf{X}_j) = I_{\text{PID}}^{(\text{syn})}(\mathbf{X}_i; \mathbf{X}_j \mid \mathbf{Y}),$$

où \mathbf{Y} peut être une troisième variable (cible à prédire) ou un contexte. Dans le cadre du DSL, il est donc pertinent d'utiliser, lorsque c'est possible, des **métriques entropiques** pour quantifier la contribution **non triviale** de chaque couple d'entités.

Une autre approche consiste à définir la synergie comme le **gain** (en termes de fonction objectif ou de performance) obtenu lorsqu'on associe deux entités, par rapport à leur utilisation séparée :

$$S(\mathcal{E}_i, \mathcal{E}_j) = \Delta(\mathcal{E}_i, \mathcal{E}_j) = \Phi(\{\mathcal{E}_i, \mathcal{E}_j\}) - [\Phi(\{\mathcal{E}_i\}) + \Phi(\{\mathcal{E}_j\})],$$

où $\Phi(\cdot)$ est un **score** ou une **mesure** de la qualité du système (ex. : taux de classification, log-likelihood, etc.). Dans ce cas :

- $S(\mathcal{E}_i, \mathcal{E}_j) > 0$ signifie qu'il y a véritablement une **valeur ajoutée** à combiner \mathcal{E}_i et \mathcal{E}_j .
- $S(\mathcal{E}_i, \mathcal{E}_j) < 0$ indique qu'il y a **inhibition** ou dégradation mutuelle.
- $S(\mathcal{E}_i, \mathcal{E}_j) = 0$ suggère une **indépendance** ou une simple addition sans synergie.

Cette formulation est souvent utilisée en pratique, car elle s'aligne directement sur un **objectif** (objectif supervisé, critère d'optimisation non supervisé, etc.). On peut de plus pondérer cette synergie par un facteur adaptatif, en tenant compte du **contexte temporel** ou **des autres entités** impliquées.

Contrairement aux approches linéaires ou statiques, le DSL prévoit que la synergie $S(\mathcal{E}_i, \mathcal{E}_j)$ soit **évolutive** au cours du temps. En d'autres termes, le réseau peut réévaluer en continu la contribution mutuelle de \mathcal{E}_i et \mathcal{E}_j . Matériellement, cela se traduit par la mise à jour d'une **pondération synergiques** $\omega_{i,j}(t)$:

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta \cdot [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \cdot \omega_{i,j}(t)],$$

où η est un taux d'apprentissage, τ un terme de régularisation (ou d'oubli). Plus la synergie entre deux entités est forte, plus leur lien s'intensifie. Au contraire, si ce lien n'apporte guère de valeur ajoutée (ou est carrément nuisible), sa pondération peut diminuer et aller jusqu'à **rompre** la connexion.

Définition Synergie binaire, ternaire, et n-aire. Dans sa version la plus simple, on considère la synergie entre **paires** d'entités $(\mathcal{E}_i, \mathcal{E}_j)$. Toutefois, nombre de scénarios exigent d'évaluer la synergie entre plusieurs entités simultanément. Dans ce cas, on généralise S à un ensemble $\{\mathcal{E}_{k_1}, \dots, \mathcal{E}_{k_m}\}$. On parle alors de **synergie n-aire**, dont la mesure n'est pas forcément la somme des synergies binaires. En effet, il se peut qu'une **triple** $\{\mathcal{E}_a, \mathcal{E}_b, \mathcal{E}_c\}$ dégage une synergie supérieure à la somme des synergies de ses paires :

$$S(\mathcal{E}_a, \mathcal{E}_b, \mathcal{E}_c) > S(\mathcal{E}_a, \mathcal{E}_b) + S(\mathcal{E}_b, \mathcal{E}_c) + S(\mathcal{E}_a, \mathcal{E}_c).$$

Ce phénomène traduit la nature **émergente** du DSL, où des ensembles d'entités peuvent coopérer de manière non triviale pour engendrer de nouvelles représentations ou actions.

Exemple scénario multimodal

Soit

- $\mathcal{E}_{\text{visuelle}}$ (extraction de caractéristiques d'une image),
- $\mathcal{E}_{\text{auditives}}$ (traits de voix, intonation),
- $\mathcal{E}_{\text{textuelles}}$ (mots-clés extraits d'une transcription).

Si $\mathcal{E}_{\text{visuelle}}$ et $\mathcal{E}_{\text{textuelle}}$ ont peu d'information en commun, leur **information mutuelle** peut être faible. Pourtant, prises ensemble, elles peuvent produire un **contexte** (ex. : “lieu de la scène + thèmes abordés verbalement”) qui aide à l'interprétation des sons (détection d'émotion). Autrement dit, c'est l'**intersection** de ces informations qui devient cruciale, expliquant une **synergie** plus forte lorsqu'on combine ces trois entités plutôt que deux à deux.

Importance pour le DSL.

Le concept de **synergie informationnelle** est ce qui **différencie** le **DSL** d'un simple système de propagation de données entre couches. Contrairement à une transmission passive, le **DSL** repose sur des interactions dynamiques où les entités cherchent activement des partenaires synergiques, ajustent leurs pondérations en fonction du niveau de synergie et forment des **clusters** ou micro-réseaux autour des synergies les plus fortes. Ce processus favorise également la fusion ou l'évolution des entités pour optimiser la coopération, un mécanisme qui sera approfondi dans les chapitres suivants.

Ainsi, la synergie agit comme un **moteur** d'auto-organisation et de **dynamique adaptative**, permettant au réseau de se **restructurer** au fil du temps, en valorisant les combinaisons d'entités les plus porteuses d'information ou de gain de performance.

1.2.3. Différence entre Interaction, Synergie et Corrélation

Lorsqu'on étudie les relations entre différentes entités d'information, il est essentiel de faire la distinction entre **interaction**, **corrélation** et **synergie**. Ces notions sont parfois utilisées de façon interchangeable, mais elles renvoient à des réalités mathématiques et conceptuelles différentes. Comprendre ces nuances permet de mieux cerner l'originalité du Deep Synergy Learning (DSL) et la portée de son concept de « synergie informationnelle ».

1.2.3.1. Interaction : une relation générique

Le terme **interaction** désigne de manière générale l'influence mutuelle que peuvent exercer deux éléments (ou plus) l'un sur l'autre. D'un point de vue mathématique, on parle souvent d'**interaction** lorsque le comportement (ou la fonction) d'un système dépend de l'état de plusieurs variables de manière **non indépendante** :

$$f(\mathbf{x}_1, \mathbf{x}_2) \neq f_1(\mathbf{x}_1) + f_2(\mathbf{x}_2).$$

Par exemple, dans un modèle statistique de type régression, l'**effet d'interaction** entre deux variables se traduit par la présence d'un terme produit $\mathbf{x}_1 \times \mathbf{x}_2$.

Une interaction ne garantit pas nécessairement un effet bénéfique ou un « plus » collectif ; elle se borne à signaler que l'état ou la valeur prise par \mathbf{x}_2 modifie l'effet de \mathbf{x}_1 (et inversement).

Ainsi, dans le DSL, de simples **interactions** peuvent exister entre des entités d'information sans pour autant impliquer une **synergie** (cette dernière requérant un effet d'émergence véritable, voir plus bas).

1.2.3.2. Corrélation : dépendance statistique (souvent linéaire)

La **corrélation** (au sens commun) mesure le degré de **dépendance statistique** entre deux variables, souvent réduit à la **corrélation linéaire** de Pearson :

$$\rho(\mathbf{X}, \mathbf{Y}) = \frac{\text{cov}(\mathbf{X}, \mathbf{Y})}{\sigma_{\mathbf{X}} \cdot \sigma_{\mathbf{Y}}}$$

avec $\text{cov}(\cdot, \cdot)$ représente la covariance, $\sigma_{\mathbf{X}}$ et $\sigma_{\mathbf{Y}}$ désignent l'écart-type de \mathbf{X} et \mathbf{Y} .

Une corrélation élevée ($\rho \approx 1$ ou $\rho \approx -1$) signifie que deux variables évoluent de façon similaire (linéairement liée), tandis qu'une corrélation nulle ($\rho \approx 0$) indique l'absence de dépendance linéaire (mais pas forcément l'absence de dépendance tout court).

Remarque : Dans un cadre non linéaire, d'autres mesures (mutual information, distance correlation, etc.) peuvent s'avérer plus pertinentes que la simple corrélation linéaire.

Corrélation \neq synergie.

Dans le cas d'une corrélation forte, on observe souvent une **redondance**. Si \mathbf{X} prédit bien \mathbf{Y} , alors \mathbf{Y} n'apporte pas nécessairement de nouvelle information. La corrélation peut aussi être trompeuse (corrélation de variables bruitées, effet de causalité inverse, variables cachées...).

Dans le DSL, deux entités très corrélées peuvent d'ailleurs être moins intéressantes (peu de gain) que deux entités faiblement corrélées, mais dont la **combinaison** génère un contenu supplémentaire.

Ainsi, une **corrélation forte** n'implique pas forcément une **synergie** ; et inversement, deux entités peuvent ne pas être corrélées mais créer, ensemble, un effet émergent.

La **synergie**, telle que définie dans le DSL, suppose un **gain** ou une **valeur ajoutée** lorsque les entités se combinent, au-delà de ce qu'elles apportent chacune de leur côté. Autrement dit :

$$S(\mathcal{E}_i, \mathcal{E}_j) > 0 \Rightarrow \text{La combinaison } \{\mathcal{E}_i, \mathcal{E}_j\} \text{ vaut plus que la somme séparée.}$$

- **Synergie \neq simple interaction** : L'interaction signale simplement une dépendance réciproque, alors que la synergie suppose qu'un nouveau niveau de fonctionnalité ou d'information émerge.

- **Synergie \neq redondance** : Deux variables très similaires (corrélées) ont peu de synergie, car prendre l'une ou l'autre n'ajoute pas grand-chose à la décision globale.
- **Synergie \neq coïncidence** : Les coïncidences peuvent être éphémères et non reproductibles. La synergie implique un **effet régulier et réel** sur l'optimisation ou la représentation interne du système.

Matériellement, dans un modèle où la fonction de coût \mathcal{L} est à minimiser (ou la fonction de performance Φ à maximiser), la synergie entre \mathcal{E}_i et \mathcal{E}_j s'exprime souvent comme :

$$\Delta_{ij} = \Phi(\{\mathcal{E}_i, \mathcal{E}_j\}) - [\Phi(\{\mathcal{E}_i\}) + \Phi(\{\mathcal{E}_j\})].$$

- Si $\Delta_{ij} > 0$, on parle de **synergie positive** (la combinaison est plus utile que la simple juxtaposition).
- Si $\Delta_{ij} < 0$, il y a **inhibition** ou **redondance néfaste**, et le couplage des entités se révèle contre-productif.
- Si $\Delta_{ij} \approx 0$, cela signifie qu'elles n'apportent pas grand-chose l'une à l'autre au regard de la tâche considérée.

Pour mieux illustrer ces différences, on peut imaginer un **diagramme** représentant trois situations :

- **Corrélation (redondance)** : Les deux entités (A et B) apportent presque la même information.
- **Interaction** : A et B se modifient mutuellement, mais sans forcément créer une nouvelle dimension.
- **Synergie** : La combinaison A + B engendre un nouveau potentiel (ex. : l'ajout d'un flux audio + flux visuel crée un contexte multimodal plus riche que n'importe lequel des flux pris isolément).

Les **cas concrets** illustrent différentes situations où la synergie joue un rôle déterminant dans l'efficacité des interactions entre entités.

Dans un **cas de redondance sans synergie**, deux capteurs de température placés au même endroit fournissent des mesures quasi identiques. Bien qu'ils soient fortement corrélés, leur apport informationnel est redondant, ce qui signifie qu'en utiliser un seul est aussi pertinent que d'en prendre deux.

Dans un **cas d'interaction sans synergie**, certaines protéines en biologie interagissent, par exemple lorsqu'une bloque l'autre, mais cette interaction ne conduit pas nécessairement à un fonctionnement plus efficace pour l'organisme. L'interaction est présente, mais elle ne génère pas de gain de performance ou d'organisation.

Dans un **cas de synergie forte**, le traitement de la parole illustre un exemple significatif. La combinaison de la lecture labiale, qui analyse les mouvements des lèvres, et du signal acoustique, qui capte l'information sonore, améliore nettement la reconnaissance par rapport à l'utilisation

isolée de l'un ou l'autre des signaux. Cet effet est particulièrement visible en environnement bruyant, où la complémentarité entre les deux sources d'information permet une meilleure robustesse du système.

Dans le **Deep Synergy Learning**, on cherche précisément à **favoriser** la création de synergies positives et à **réduire** (voire éliminer) les liens qui relèvent de la simple redondance ou d'interactions stériles. Les règles de mise à jour des **pondérations synergiques** (détaillées dans les chapitres dédiés) sont conçues pour ajuster dynamiquement les connexions entre entités en fonction de leur synergie. Elles permettent de **renforcer** les liens entre entités lorsque $\Delta_{ij} > 0$, favorisant ainsi les interactions bénéfiques. À l'inverse, elles visent à **diminuer** ou rompre les connexions lorsque la corrélation entre entités ne procure aucun gain ou engendre un effet négatif, c'est-à-dire lorsque $\Delta_{ij} < 0$. Enfin, ces règles facilitent la **détection** de synergies **n-aires**, où plusieurs entités interagissent pour former des **micro-réseaux auto-organisés**, optimisant ainsi la structure adaptative du **SCN**.

Cette démarche permet d'éviter l'explosion combinatoire (en évaluant toutes les combinaisons) grâce à un **mécanisme dynamique** où les liens se forment ou se défont au fil de l'apprentissage, suivant les feedbacks de performance ou des indicateurs entropiques.

Dans le DSL, c'est précisément cette notion de synergie, mesurée et mise à jour en continu, qui permet de construire des **clusters** d'entités coopératives, de faire émerger de **nouvelles représentations**, et de potentialiser la **résilience** du système face à la variabilité des données. Les sections ultérieures reviendront sur la façon dont ce mécanisme de synergie se met en place dans une **approche auto-organisée** (1.2.4) et comment il se distingue des **réseaux neuronaux** traditionnels (1.2.5).

1.2.4. Approche Hiérarchique vs Approche Auto-Organisée

Les approches hiérarchiques traditionnelles, largement utilisées dans le Deep Learning, reposent sur la succession de **couches** (layers) qui transforment progressivement les données d'entrée jusqu'à aboutir à une sortie (une prédiction, une classification, etc.). À chaque couche, on opère une **composition** de fonctions (le plus souvent linéaires, suivies de non-linéarités comme ReLU ou sigmoid). Par opposition, l'**approche auto-organisée** (telle qu'on la retrouve dans le Deep Synergy Learning, DSL) met l'accent sur la **capacité du réseau à reconfigurer** ou **réorganiser** sa structure interne en fonction des synergies détectées, plutôt que de s'en tenir à une architecture rigide et prédéfinie.

Dans cette section, nous allons approfondir cette différence en examinant les principes fondamentaux de l'approche hiérarchique, pour mieux comprendre comment le DSL, en tant qu'approche auto-organisée, propose une alternative à la fois plus **dynamique** et plus **adaptative**.

1.2.4.1. Fondements de l'approche hiérarchique

Dans un **réseau hiérarchique** traditionnel (tel qu'un réseau de neurones profond), le **principe du traitement en cascade** impose aux données **x** de traverser une succession de transformations linéaires et non linéaires :

$$\mathbf{h}^{(1)} = f^{(1)}(\mathbf{x}), \quad \mathbf{h}^{(2)} = f^{(2)}(\mathbf{h}^{(1)}), \quad \dots, \quad \mathbf{h}^{(L)} = f^{(L)}(\mathbf{h}^{(L-1)}),$$

où $\mathbf{h}^{(\ell)}$ représente la “représentation cachée” extraite à la couche ℓ , et $f^{(\ell)}$ désigne un opérateur paramétrique (incluant poids et fonction d’activation). Dans ce schéma, l’information **circule** essentiellement **de bas en haut**, et les éventuelles boucles de rétroaction (feedback top-down) demeurent limitées ou spécialisées, comme dans les architectures RNN ou LSTM.

Cette organisation induit une **séparation des rôles** :

- Les **premières couches** traitent des descripteurs “bas niveau” (par exemple, repérer des bords pour une image, des phonèmes pour un signal audio, etc.).
- Les **couches intermédiaires** approfondissent la combinaison de ces descripteurs, extrayant des motifs plus complexes.
- Les **dernières couches** produisent la **décision finale** (classe, score, etc.).

Toutefois, cette hiérarchie s’accompagne d’une **certaine rigidité**. Une fois le nombre de couches, la taille de chacune, et la nature des connexions (dense, convolutionnelle, récurrente...) choisis, la **structure** du réseau reste figée pendant la phase d’apprentissage. Seuls les **poids** sont ajustés par descente de gradient ou par l’une de ses variantes, tandis que la topologie globale demeure invariable.

1.2.4.2. Limites de l’approche hiérarchique

Les **réseaux hiérarchiques** classiques présentent plusieurs limites marquantes. Ils restent fortement **dépendants** à la supervision, nécessitant de grandes quantités de données annotées pour ajuster leurs poids internes. À mesure que leur architecture croît, une **prolifération exponentielle** du nombre de paramètres alourdit le coût en calcul et en mémoire, tout en rendant le réseau plus difficile à **interpréter** et à **déboguer**.

Ces réseaux souffrent également d’un **manque d’adaptabilité**. Lorsqu’un **domain shift** survient, c’est-à-dire un changement dans la distribution des données, il devient nécessaire de procéder à un réapprentissage ou à un ajustement considérable, faute de mécanismes permettant une **reconfiguration dynamique**. Enfin, leur **faible modularité** constitue une autre contrainte. Même si les couches s’empilent, elles ne peuvent pas **échanger librement** en dehors des cheminements rigides imposés par l’architecture.

En résumé, un réseau hiérarchique fonctionne efficacement dans des contextes précis, lorsque de grands volumes de données annotées sont disponibles, mais il ne propose pas de **remodelage** structurel en cours d’apprentissage ni de recours proactif à des **synergies** entre sources d’information hétérogènes.

1.2.4.3. Principes de l'approche auto-organisée

Dans une **approche auto-organisée**, caractéristique du **DSL**, la **plasticité topologique** joue un rôle central. Plutôt que de reposer sur des couches fixes, les **entités d'information** peuvent spontanément **former des clusters**, **établir ou rompre des liaisons**, et même **fusionner** lorsque leur synergie atteint un niveau suffisant (cf. sections 1.2.2 et 1.2.3). L'architecture n'est donc plus imposée de manière rigide, mais se construit et se reconstruit au fil de l'apprentissage, s'adaptant en fonction des besoins et de l'évolution des données.

Cette **évolution dynamique** repose sur l'adaptation continue des pondérations synergiques $\omega_{i,j}$ suivant une règle telle que

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)],$$

où η est un taux d'apprentissage, τ un coefficient de régularisation et $S(\mathcal{E}_i, \mathcal{E}_j)$ la synergie entre les entités \mathcal{E}_i et \mathcal{E}_j . Dans ce cadre, une synergie **positive** renforce la connexion, tandis qu'une synergie **négative** l'affaiblit. Ainsi, les entités dont la coopération est bénéfique sont encouragées à établir (ou consolider) leurs liens, tandis que les connexions moins productives s'éteignent naturellement.

Au fil de ce processus d'ajustement, des **micro-réseaux** ou **clusters auto-organisés** émergent dès lors que les synergies mutuelles s'élèvent entre certaines entités. Ces agrégats peuvent apparaître, se scinder ou disparaître, reflétant les évolutions des données ou l'arrivée d'interactions nouvelles. Par ailleurs, cette liberté structurelle favorise une **coopération multi-flux**, en autorisant les entités visuelles, auditives, textuelles, etc. à s'influencer **directement**, sans passer par un chemin prédéfini de "couches". Par exemple, une entité auditive peut détecter à l'instant t une forte synergie avec une entité visuelle et, de ce fait, former un cluster pour la durée nécessaire ; plus tard, si les conditions changent, elle peut s'éloigner de ce groupe pour établir d'autres coopérations plus pertinentes.

1.2.4.4. Modélisation mathématique d'une architecture auto-organisée

Dans une **architecture auto-organisée**, on ne définit plus un enchaînement linéaire $\mathbf{h}^{(1)} \rightarrow \mathbf{h}^{(2)} \rightarrow \dots$; à la place, on conçoit un **graphe** $G(t)$ dont les **nœuds** sont les entités $\{\mathcal{E}_1, \dots, \mathcal{E}_n\}$ et dont les **arêtes** représentent les pondérations synergiques $\omega_{i,j}(t)$. L'évolution du réseau se décrit alors par une fonction de mise à jour :

$$G(t+1) = \mathcal{U}[G(t), S(\cdot, \cdot)],$$

où \mathcal{U} tient compte des **critères de synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$. Le système prend ainsi la forme d'un **Système Dynamique Non Linéaire**, se réorganisant de manière à privilégier les **combinatoires** d'entités jugées les plus utiles ou performantes.

Il est fréquent d'ajouter un mécanisme d'**énergie libre** ou de **coût global** :

$$\mathcal{J}(G) = - \sum_{i,j} \omega_{i,j} S(\mathcal{E}_i, \mathcal{E}_j) + \alpha \|\boldsymbol{\omega}\|^2,$$

de façon à **régulariser** la taille du réseau et à éviter que le nombre de connexions ne “flambe” de manière excessive. Dans ce cadre, la mise à jour peut s’effectuer via une **descente de gradient** (ou un algorithme d’optimisation inspiré des systèmes complexes, comme un algorithme génétique ou un recuit simulé), conduisant progressivement à une **organisation** qui valorise les synergies tout en restreignant les liaisons redondantes.

1.2.4.5. Avantages et défis de l’auto-organisation

Dans une **approche auto-organisée**, l’**adaptabilité** constitue un atout majeur. Le réseau peut s’ajuster en continu face à l’arrivée de nouvelles données ou à un changement de distribution sans qu’il soit nécessaire de repenser entièrement son architecture. La **multi-modalité** y est gérée de manière native. Les entités issues de différentes sources comme l’audio, l’image ou le texte peuvent **s’influencer mutuellement** et **former des clusters** multimodaux. Cette capacité favorise la **fusion** des flux d’informations et l’exploitation de leurs synergies. Cette dynamique ouverte introduit également un **potentiel créatif**. L’**émergence** de combinaisons inédites entre entités permet de révéler des **patterns** jusque-là invisibles et difficilement capturables par des architectures hiérarchiques classiques.

Toutefois cet avantage s’accompagne de plusieurs **défis** majeurs. La **complexité de contrôle** peut devenir un obstacle si aucun mécanisme de régulation comme le facteur τ ou des pénalités n’est mis en place. Un excès de connexions ou des boucles auto-renforcées peuvent entraîner des oscillations. Le **coût de calcul** représente également un défi. L’évaluation de la synergie entre un grand nombre d’entités nécessite l’usage d’heuristiques ou de méthodes parcimonieuses pour rester viable en pratique. L’**interprétabilité** peut aussi poser problème. Même si l’auto-organisation tend à faire émerger des clusters plus significatifs, l’évolution dynamique et permanente de la structure complique l’analyse détaillée du fonctionnement interne du réseau.

1.2.4.6. Comparaison synthétique

| Caractéristique | Approche Hiérarchique | Approche Auto-Organisée (DSL) |
|---------------------------------|---------------------------------|-------------------------------------------|
| Architecture | Fixe (définie a priori) | Flexible (graphe évolutif) |
| Propagation de l’info | Principalement feed-forward | Libre (coopération directe entre entités) |
| Formation des connexions | Statique (paramètres ajustés) | Dynamique (création / rupture de liens) |
| Apprentissage | Descente de gradient classique | Mise à jour des synergies (pondérations) |
| Multimodalité | Fusion tardive (généralement) | Intégration native, clusters multimodaux |
| Adaptation continue | Limité (fine-tuning, transfert) | Fort (reconfiguration à la volée) |
| Exemples | CNN, RNN, Transformers | Synergistic Connection Network (SCN) |

Les sections suivantes (1.2.5 et suivantes) reviendront sur la comparaison plus directe entre les **réseaux neuronaux** traditionnels et les **réseaux synergiques**, tout en introduisant la terminologie spécifique (1.2.6) et des **exemples** (1.2.7) illustrant la pertinence de l’auto-organisation dans différents contextes naturels.

1.2.5. Réseaux Neuronaux Traditionnels vs Réseaux Synergiques

Dans les sections précédentes, nous avons présenté les notions de **synergie informationnelle**, de **corrélation** et de **plasticité** structurelle dans une approche **auto-organisée**. Il est maintenant temps de faire un **parallèle** entre, d’une part, les **réseaux neuronaux profonds** (ou traditionnels) tels qu’on les connaît en apprentissage profond (Deep Learning) et, d’autre part, les **réseaux synergiques** comme envisagés dans le DSL (Deep Synergy Learning). Cette comparaison aidera à mettre en lumière ce que le DSL apporte de différent par rapport aux architectures classiques (CNN, RNN, Transformers, etc.).

1.2.5.1. Structure et dynamique d’apprentissage

Dans les **réseaux neuronaux traditionnels**, la structure est définie dès le départ. Le nombre de couches et leur type, qu’elles soient convolutionnelles, récurrentes ou fully-connected, sont déterminés à l’avance ainsi que l’agencement de ces couches et la manière dont elles s’enchaînent. Chaque couche s’appuie sur la représentation produite par la précédente, imposant une **hiérarchie** explicite. L’apprentissage se fait en ajustant les poids et les biais via une descente de gradient ou l’une de ses variantes (Adam, RMSProp, etc.), tandis que la **propagation de l’information** suit principalement un chemin **feed-forward** ; même lorsque des boucles internes existent (RNN, LSTM), elles demeurent cantonnées à la topologie imposée.

À l’inverse, dans les **réseaux synergiques** (au sein du DSL), la topologie se veut **flexible et évolutive** : on ne parle plus de “couches” stricto sensu, mais d’un **ensemble d’entités d’information** liées entre elles par des pondérations **synergiques** (voir 1.2.2 et 1.2.4). Ces pondérations ne sont pas seulement ajustées ; elles peuvent aussi être créées, renforcées ou rompues, selon la synergie détectée. Cette possibilité de **reconfiguration** permanente marque la différence. Le réseau ne se contente pas d’empiler des couches mais s’**auto-organise** en **clusters** lorsque la synergie l’exige. De plus, l’information ne circule pas selon une progression linéaire ; elle peut **transiter** entre toutes les entités jugées “synergiques”, adoptant ainsi une **approche distribuée** plus proche d’un écosystème vivant que d’un pipeline hiérarchisé.

1.2.5.2. Mesure de la performance et critères d’apprentissage

Dans les **objectifs classiques** du Deep Learning, on minimise une **fonction de coût** $\mathcal{L}(\theta)$ (par exemple, l’entropie croisée ou la MSE), à l’aide d’une **backpropagation** qui calcule les gradients. Les performances sont ensuite mesurées selon des métriques comme la **précision**, le **rappel** ou le **F1-score**, en fonction du type de tâche (classification, régression, etc.).

En revanche, dans une **approche DSL**, ces **objectifs traditionnels** (par exemple, la précision en classification) coexistent avec des **fonctions de synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$ qui orientent l'apprentissage. Les mises à jour des pondérations synergiques $\omega_{i,j}$ tiennent compte de ces scores de synergie, **favorisant** les liaisons dont la synergie s'avère significative, tout en permettant la **création**, le **renforcement** ou la **dissolution** de connexions. On peut également définir une **fonction globale** $J(G)$, laquelle agrège la **somme** (ou autre forme d'agrégation) des synergies et **pénalise** la multiplication de connexions redondantes. Le réseau se comporte alors comme un **système dynamique**, visant à concilier la **minimisation** d'une perte liée aux tâches classiques (p. ex. une fonction de classification) et la **maximisation** de la synergie informationnelle au sens large.

1.2.5.3. Comparaison de la gestion de la multimodalité

Dans les **réseaux neuronaux traditionnels**, la **multimodalité** est gérée en définissant à l'avance des voies de traitement spécifiques pour l'audio, l'image ou le texte. Ces différentes modalités sont ensuite **fusionnées** à un certain stade du réseau, soit au niveau de couches intermédiaires, soit plus souvent par une fusion tardive vers la fin du pipeline. L'architecture doit donc être **explicitement conçue** pour chaque canal (par exemple, un CNN dédié à l'image, un RNN ou un Transformer pour le texte, puis un module spécialisé pour agréger les différents flux). La synergie potentielle entre ces canaux se découvre **indirectement**, via la backpropagation, mais la structure globale du réseau — et la manière dont les canaux se croisent — reste imposée de l'extérieur.

À l'inverse, dans les **réseaux synergiques** du DSL, les **entités** associées à divers canaux (audio, visuel, textuel, etc.) ont la possibilité de **se “découvrir” mutuellement** au fil de l'apprentissage. Si l'audio et l'image présentent une forte synergie, elles peuvent **former un cluster multimodal** de manière autonome, sans qu'une couche de fusion spécifique ne soit paramétrée au préalable. Cette approche facilite la **co-évolution** des représentations. Si le flux visuel est perturbé par du bruit à un moment donné, l'entité visuelle peut s'appuyer davantage sur les canaux texte ou audio, à condition qu'une synergie élevée soit détectée. Ainsi, la multimodalité se développe de façon **organique**, guidée par la recherche de gains effectifs de performance ou d'information.

1.2.5.4. Rôle de la rétropropagation et alternatives

Dans les **réseaux neuronaux profonds** classiques, la **rétropropagation** constitue le mécanisme standard pour ajuster les **poids** de couche en couche ; on dispose alors d'une **architecture** explicitement définie de bout en bout et d'un **objectif** scalaire unique qui oriente la descente de gradient (p. ex. l'entropie croisée). En revanche, dans un **DSL**, les **mises à jour** des pondérations synergiques se réalisent de manière **distribuée**, souvent selon des règles plus locales (inspirées, par exemple, d'approches “Hebbiennes généralisées” ou d'une évaluation directe des gains de performance obtenus). Un **objectif global** peut persister (comme un taux de reconnaissance), mais la **découverte** de synergies s'opère fréquemment dans des configurations plus indépendantes du gradient global.

La **rétropropagation** n'est pas pour autant **exclue**. Il est possible d'adopter un **système hybride** où la backpropagation s'applique à certains sous-modules, tandis que la **formation** et la **reconfiguration** du graphe synergique suivent des lois d'auto-organisation distinctes. Cette

approche offre une **flexibilité** accrue. Certaines parties du réseau continuent d'être ajustées finement par le gradient, tandis que le réseau conserve la capacité de **découvrir** et de **renforcer** localement des liaisons synergiques au-delà des contraintes d'une architecture fixe.

1.2.5.5. Robustesse et adaptation continue

Dans les **réseaux neuronaux traditionnels**, la **robustesse** dépend essentiellement de la qualité du jeu d'entraînement et de plusieurs mécanismes de **régularisation** (dropout, batch normalization, etc.). L'**adaptation** à un nouveau domaine s'effectue souvent par un transfert d'apprentissage (transfer learning), suivi d'un **fine-tuning** partiel ou complet des poids du réseau. Cependant, ce procédé peut exposer le système au risque de **catastrophic forgetting** lorsqu'il est utilisé pour apprendre en continu une succession de tâches. Les poids ajustés pour les nouvelles données ont alors tendance à effacer les connaissances acquises précédemment.

Dans les **réseaux synergiques**, au contraire, la **structure évolutive** permet au système d'**allouer** de nouvelles entités ou de **renforcer** certains liens pour absorber plus facilement un **changement** de données, sans exiger un réapprentissage complet de l'ensemble du réseau. Les **clusters** déjà formés pour des tâches précédentes peuvent coexister dans la nouvelle configuration, au lieu d'être remplacés ou écrasés. Ainsi, un réseau synergique peut mieux **retenir** l'expérience passée (réduisant d'autant le **catastrophic forgetting**) et faire preuve d'une plus grande **flexibilité** quand son environnement ou sa mission évoluent.

1.2.5.6. Réseaux traditionnels et synergiques : cohabitation possible ?

Les **réseaux synergiques** n'ont pas vocation à **remplacer** purement et simplement les architectures neuronales traditionnelles. Au contraire, divers **scénarios de cohabitation** sont envisageables. On peut, par exemple, adopter une **approche hybride**, dans laquelle un pipeline CNN (pour l'image) ou Transformer (pour le texte) extrait des **représentations** initiales ; ces représentations deviennent ensuite des **entités** au sein d'un réseau synergique, lequel peut alors coopérer et se reconfigurer de manière plus libre.

Dans certains **systèmes complexes**, on peut aussi instaurer une **transition progressive**, en commençant par des couches de feature extraction classiques, puis en insérant une **couche synergique** à un stade où les divers canaux se croisent. De cette façon, on préserve la puissance des modèles traditionnels pour l'extraction de caractéristiques tout en intégrant la logique auto-organisée et adaptative du DSL à un niveau plus élevé.

Enfin, il est possible de développer des **extensions spécialisées**, par exemple un composant auto-organisé dédié à la **fusion multimodale** ou à la **gestion de multiples contextes**, tandis que la classification finale demeure assurée par un réseau fully-connected ordinaire. L'essentiel est d'exploiter la **flexibilité** des réseaux synergiques dans les domaines où ils excellent — par exemple, l'émergence dynamique de clusters — tout en s'appuyant sur l'expérience accumulée des architectures neuronales traditionnelles.

1.2.5.7. Synthèse et perspectives

| Aspect | Réseaux Neuronaux Traditionnels | Réseaux Synergiques (DSL) |
|-----------------------------|------------------------------------------------------|-------------------------------------------------------------|
| Topologie | Fixe, pré-spécifiée | Évolutive, auto-organisée |
| Propagation | Hiérarchique, feed-forward | Dispersée, multidirectionnelle |
| Apprentissage | Backpropagation end-to-end | Règles locales + mise à jour synergie |
| Évolution temporelle | Nécessite du re-training pour s'adapter | Adaptation dynamique à la volée |
| Gestion multimodale | Fusion tardive ou intermédiaire, souvent manuelle | Fusion spontanée via synergie et création de clusters |
| Robustesse | Vulnérabilité à l'overfitting, besoin de régulariser | Auto-régulation via le feedback de synergie |
| Applications | Classification, régression, vision, NLP... | Idem, mais avec en plus la souplesse et l'auto-organisation |

En conclusion, les **réseaux neuronaux traditionnels** et les **réseaux synergiques** diffèrent principalement par la **structure**, la **dynamique d'apprentissage** et la **capacité d'auto-organisation**. Le **Deep Synergy Learning** apporte une philosophie plus **organique**, inspirée des systèmes complexes, pour que l'intelligence artificielle puisse gérer la **variabilité**, la **multimodalité**, et s'auto-adapter en continu.

La section suivante (1.2.6) clarifiera la **terminologie** propre au DSL — notamment les notions de **clusters**, **entités**, **pondérations synergiques**, etc. — puis nous verrons (1.2.7) des **exemples illustratifs** tirés de la nature ou d'applications concrètes, afin de matérialiser les principes évoqués dans ce chapitre.

1.2.6. Terminologies Employées dans le DSL

Au fil des sections précédentes, plusieurs notions-clés sont apparues pour décrire les principes du **Deep Synergy Learning (DSL)**. Il est important de les clarifier et de les organiser en un vocabulaire cohérent, car ces termes forment la **boîte à outils conceptuelle** indispensable pour aborder les mécanismes internes et les applications pratiques du DSL. Dans cette section, nous passons en revue les principaux termes et leur signification, en soulignant les liens entre eux.

1.2.6.1. Entité d'Information (ou “Information Entity”)

Définition. Dans le DSL, une **entité d'information** (souvent notée \mathcal{E}_i) représente l'unité fondamentale du système. Contrairement à un simple vecteur de données, une entité est un *objet d'apprentissage* pouvant inclure :

- Une **représentation** (par ex. un vecteur, un tenseur, ou même une distribution).
- Des **paramètres internes** (θ_i) et un **état** ($\mathbf{s}_i(t)$).

- Un **historique** (ou “mémoire”) de ses interactions passées.

C’est à travers ces entités que s’établissent les **synergies** et que se construit la dynamique de l’apprentissage. En pratique, toute source d’information (une image, un signal audio, un embedding textuel, etc.) peut être encapsulée sous forme d’entité.

1.2.6.2. Synergie (ou “Synergy”)

Définition. La **synergie** entre deux (ou plusieurs) entités est la mesure de la **valeur ajoutée** qu’elles obtiennent en coopérant, par rapport à ce qu’elles pourraient réaliser indépendamment (voir 1.2.2). Elle se note souvent $S(\mathcal{E}_i, \mathcal{E}_j)$ pour les paires, et peut être généralisée à des ensembles $\{\mathcal{E}_{k_1}, \dots, \mathcal{E}_{k_m}\}$.

Formes de mesure.

- **Informationnelle** : Basée sur l’entropie, l’information mutuelle, ou d’autres métriques de la théorie de l’information.
- **Basée sur la performance** : Différence de score (classification, regression, etc.) quand on associe \mathcal{E}_i et \mathcal{E}_j .
- **Hybride** : Combinaison d’un critère d’information et d’un critère de performance.

La synergie est la **force motrice** du **DSL**. Elle oriente la création, la rupture ou le renforcement des connexions entre entités, comme détaillé dans la section sur les **pondérations synergiques**.

1.2.6.3. Pondérations Synergiques (ou “Synergistic Weights”)

Définition. Les **Pondérations Synergiques** notées $\omega_{i,j}(t)$, ce sont les **coefficients** qui caractérisent la relation dynamique entre deux entités \mathcal{E}_i et \mathcal{E}_j à l’instant t .

Souvent modélisée par une équation du type

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)],$$

où η est un taux d’apprentissage, τ un terme de régularisation, et $S(\mathcal{E}_i, \mathcal{E}_j)$ la synergie entre les entités.

Les pondérations synergiques constituent la **matrice d’adjacence** d’un *graphe évolutif* :

$$W(t) = [\omega_{i,j}(t)]_{i,j}.$$

Elles déterminent quelles entités sont fortement liées (hautes synergies) et lesquelles le sont moins voire pas du tout (synergie quasi nulle).

1.2.6.4. Cluster (ou “Micro-Réseau”)

Définition. Un *cluster* est un **sous-ensemble** d’entités qui présentent entre elles une synergie élevée, formant ainsi une structure cohérente et **auto-organisée**.

Les entités $\{\mathcal{E}_1, \dots, \mathcal{E}_k\}$ tendent à se regrouper si leurs **pondérations synergiques** mutuelles sont supérieures à un certain seuil θ , ou si elles maximisent un critère global (p. ex. somme des synergies internes au cluster).

Les **clusters synergiques** formés au sein d’un DSL jouent un **rôle** fondamental à deux niveaux. Ils **favorisent** d’abord la **coopération locale** en permettant aux entités d’un même cluster d’échanger de manière intensive. Chaque entité apporte ses données ou compétences spécifiques, renforçant ainsi la synergie collective.

Ils **facilitent** ensuite l’**adaptation** du réseau. Ces clusters peuvent fusionner pour mieux gérer de nouveaux contextes lorsque leur compatibilité est élevée, ou se scinder lorsqu’un manque de synergie interne apparaît.

Grâce à ce double mécanisme combinant **coopération accrue** et **flexibilité structurelle**, le système reste **résilient** et capable d’évoluer face aux changements de tâches ou d’environnements.

1.2.6.5. Synergistic Connection Network (SCN)

Définition. Le SCN représente l’**infrastructure** du DSL, c’est un **réseau** dont les *nœuds* sont les entités $\{\mathcal{E}_i\}$ et dont les *arêtes* sont les pondérations $\{\omega_{i,j}\}$.

Contrairement à un réseau de neurones statique, le SCN est **dynamique**. Au fil du temps ou des itérations d’apprentissage, de nouvelles connexions apparaissent, tandis que d’autres se suppriment ou s’affaiblissent, favorisant ainsi l’émergence de **clusters** adaptatifs.

L’**objectif** central du SCN consiste à **exploiter** les **synergies** entre entités de manière à *auto-organiser* le flot d’information et, ce faisant, à optimiser la **performance** globale du système, qu’il s’agisse d’une tâche supervisée ou non supervisée. L’idée est de permettre aux liens synergiques les plus pertinents de se renforcer, afin que le réseau dirige spontanément les informations vers les chemins les plus efficaces. Ainsi, l’architecture se réagence en fonction des besoins (ou des données) pour offrir un apprentissage et un traitement des informations plus rapide et plus robuste, sans nécessiter de contrôle externe permanent.

1.2.6.6. Auto-Organisation

Définition. **Auto-organisation** désigne la **capacité** d’un réseau à *se structurer* et *se reconfigurer* de façon autonome, sans intervention ou contrôle direct de l’extérieur (cf. section 1.2.4). Ce phénomène repose sur une **évaluation continue** de la synergie entre l’ensemble ou une partie des entités. À chaque itération, les pondérations $\omega_{i,j}$ sont **mises à jour** selon une règle d’adaptation, tandis que des **clusters** peuvent se **former** ou se **dissoudre** en fonction des tendances observées.

L'**objectif** de ce mécanisme est triple. D'abord, il s'agit d'acquérir une **robustesse** accrue face aux perturbations, car le réseau peut se réorganiser spontanément lorsque des défaillances ou des changements surviennent. Ensuite, cette approche permet de gérer naturellement la **multimodalité**. Plutôt que de cloisonner les entités visuelles, auditives ou autres, elle leur permet de s'associer ou de se séparer en fonction de leurs synergies. Enfin, l'auto-organisation ouvre la voie à un **apprentissage continu**, dans lequel de nouvelles représentations émergentes se forment au fil du temps, sans imposer la rigidité d'un schéma hiérarchique figé.

1.2.6.7. État (ou “State”) d’une Entité

Définition. Chaque entité \mathcal{E}_i dispose d'un **état interne** $\mathbf{s}_i(t)$, souvent représenté par un vecteur de dimension d , qui synthétise son “histoire” ou son “contexte” à l'instant t . Cet état évolue selon une **fonction d'actualisation** F , de la forme

$$\mathbf{s}_i(t+1) = F(\mathbf{s}_i(t), \{\omega_{i,j}(t)\}_j, \dots),$$

inspirée, par exemple, de modèles dynamiques ou de mécanismes de type RNN ou “Hebb étendu”. Le **rôle** de $\mathbf{s}_i(t)$ est déterminant pour la réactivité de l'entité. Une entité ayant déjà établi de fortes coopérations avec une autre \mathcal{E}_j est généralement plus encline à **se synchroniser** de nouveau avec elle, la mémoire de leurs interactions passées augmentant la probabilité d'une synergie future.

1.2.6.8. Mécanismes de Fusion et de Dissociation

Deux (ou plusieurs) entités $\{\mathcal{E}_i, \mathcal{E}_j, \dots\}$ peuvent **fusionner** s'il s'avère qu'elles sont presque systématiquement dans un même cluster et qu'elles partagent une forte synergie dans la durée. Cette fusion se modélise par la création d'une **nouvelle entité** $\mathcal{E}_{\text{fusion}}$, qui combine leurs états, leurs mémoires et leurs représentations.

Lorsqu'une entité \mathcal{E}_k se trouve dans un cluster peu cohérent (synergie moyenne ou négative), elle peut se **retirer** du cluster ou rompre une fusion antérieure.

Ces mécanismes confèrent au DSL une **plasticité structurale** comparable à celle de certains systèmes biologiques (cerveau, colonies d'insectes, etc.), favorisant l'adaptation face à de nouveaux contextes ou de nouvelles tâches.

1.2.6.9. Énergie ou Fonction Globale \mathcal{J}

On peut parfois définir une **fonction** $\mathcal{J}(G)$ — parfois appelée “énergie libre” ou “coût global” — qui regroupe, d'une part, les **synergies** positives entre les entités et, d'autre part, un **terme** de pénalisation destiné à éviter une **surabondance** de connexions. Par exemple :

$$\mathcal{J}(G) = - \sum_{i,j} \omega_{i,j} S(\mathcal{E}_i, \mathcal{E}_j) + \alpha \|\boldsymbol{\omega}\|^2,$$

où ω désigne le vecteur de toutes les pondérations synergiques, et α un coefficient de régulation. **Minimiser** $\mathcal{J}(G)$ revient alors à **maximiser** la somme de synergies utiles tout en **limitant** la prolifération de liens non pertinents. Il s’agit ainsi d’une démarche **globale** pour piloter l’auto-organisation du réseau, puisqu’elle encourage les connexions réellement productives tout en imposant un frein à celles qui n’apporteraient aucun gain substantiel.

1.2.6.10. Apprentissage Continu (ou “Lifelong Learning”)

Dans le **DSL**, l’apprentissage ne se limite pas à une **phase offline** unique ; le réseau peut, au contraire, **évoluer** continuellement face à un flux de données **online**, en réajustant de façon permanente les pondérations $\omega_{i,j}(t)$ ainsi que la configuration des clusters.

Cet **apprentissage continu** présente plusieurs **avantages**. D’une part, il offre une **tolérance** accrue aux perturbations comme le bruit, les changements dans la distribution des données ou l’apparition de nouvelles classes et contextes. Cette capacité permet au système de s’adapter plus facilement à des environnements non stationnaires. D’autre part, il contribue à la **réduction** du phénomène de “forgotten knowledge”, puisque les clusters formés pour des tâches antérieures peuvent être préservés et ainsi servir de base à des transferts de connaissances ultérieurs.

1.2.6.11. Terminologies récurrentes (Synthèse)

Pour faciliter la lecture et l’implémentation, voici un **récapitulatif** des principales terminologies :

1. \mathcal{E}_i : Entité d’information numéro i .
2. $S(\mathcal{E}_i, \mathcal{E}_j)$: Synergie entre entités i et j .
3. $\omega_{i,j}(t)$: Pondération synergique (ou lien) entre entités i et j à l’instant t .
4. $W(t) = [\omega_{i,j}(t)]$: Matrice (ou graphe) des pondérations synergiques.
5. **Cluster** : Sous-groupe d’entités fortement liées (hautes $\omega_{i,j}$).
6. **SCN** : *Synergistic Connection Network*, la structure dynamique qui évolue selon les lois d’adaptation.
7. **Auto-organisation** : Processus par lequel la structure $W(t)$ se réarrange spontanément.
8. **Fonction \mathcal{J}** : Mesure globale de la qualité ou de l’état du réseau (peut inclure la somme des synergies, des pénalités, etc.).

1.2.6.12. Conclusion

Ces termes comme **entité**, **synergie**, **pondération synergique**, **cluster**, **SCN** et **auto-organisation** constituent le *lexique de base* du **DSL**. Chaque concept est interdépendant. Les **entités** interagissent

à travers des **pondérations synergiques** qui façonnent le SCN. Celui-ci se **réorganise en clusters** selon un mécanisme d'**auto-organisation** guidé par la **synergie** et, éventuellement, par une **fonction globale** J .

Dans la section suivante (1.2.7), nous illustrerons ces principes par des **exemples concrets**, qu'ils proviennent de la nature (inspirations biologiques) ou d'applications pratiques (cas d'études multimodales, émergence de schémas cognitifs, etc.). Ce sera l'occasion de vérifier comment l'utilisation rigoureuse de cette terminologie peut clarifier la **logique** et la **mise en œuvre** du Deep Synergy Learning.

1.2.7. Exemples Illustratifs de la Synergie dans la Nature

Les principes de **synergie informationnelle** et d'**auto-organisation** que promeut le Deep Synergy Learning (DSL) trouvent de nombreux échos dans les systèmes naturels. Qu'il s'agisse de colonies d'insectes, de réseaux neuronaux biologiques, d'écosystèmes ou de synchronisations collectives, on observe des processus où l'**ensemble** dépasse la **somme de ses parties**, grâce à des mécanismes coopératifs distribués. Les sous-sections suivantes illustrent comment ces phénomènes naturels inspirent l'approche synergiques du DSL.

Colonies d'Insectes et Intelligence Collective

Les colonies de fourmis, d'abeilles ou de termites sont des exemples emblématiques d'**intelligence collective**. Chaque individu, malgré des capacités limitées, contribue à la réalisation de tâches complexes comme la construction de nids sophistiqués ou l'optimisation de la recherche de nourriture. Ces interactions reposent sur des **signaux locaux** (phéromones, contacts antennaires, etc.), sans qu'aucune entité centrale ne dirige l'ensemble. L'émergence d'une organisation globale, comme le traçage de pistes ou la réparation du nid, résulte donc d'une **coopération** distribuée entre entités locales.

Dans le **Deep Synergy Learning (DSL)**, les **entités d'information** fonctionnent de manière analogue aux insectes sociaux. Elles établissent ou rompent des liens en fonction de la **pertinence** ou de la **synergie** qu'elles y perçoivent. Ces **connexions synergiques** évoluent en permanence, à l'image des fourmis qui renforcent ou abandonnent certains chemins selon leur utilité.

La formation de **clusters** d'entités synergiques dans le **DSL** rappelle les **micro-sociétés** au sein d'une colonie d'insectes, où chaque groupe se spécialise dans une tâche particulière. Cette **auto-organisation** spontanée illustre la puissance d'un système distribué. Sans planification rigide, l'ensemble du réseau s'adapte et se coordonne pour atteindre un objectif global.

Synergies dans le Cerveau et les Réseaux Neuronaux Biologiques

Dans le **cerveau** humain ou animal, la **plasticité synaptique** illustre la puissance d'un réseau extrêmement **connecté**, dans lequel les synapses s'ajustent en fonction des interactions locales. Lorsque deux neurones s'associent régulièrement pour traiter un même stimulus, leur **synapse** se

renforce (potentialisation à long terme) ; ce phénomène rappelle la **mise à jour** des liaisons synergiques dans le DSL, où les liens forts se consolident à mesure que les entités coopèrent efficacement.

Les neurosciences démontrent également la formation d'**assemblées neuronales** associées à un concept ou à un stimulus précis. Ces assemblées se créent ou se dissolvent selon le **contexte** ou la **tâche** du moment. De la même manière, le DSL autorise la **création** et la **dissolution de clusters** d'entités d'information, la synergie entre ces entités évoluant dans le temps pour s'adapter aux besoins et aux données.

Enfin, ce parallèle s'étend à la **mémoire** et à l'**apprentissage**. Dans le cerveau, les synapses permettent à un neurone de conserver les connexions consolidées lors d'apprentissages antérieurs, lui offrant ainsi la capacité de "se souvenir" des interactions passées. Dans le DSL, chaque entité conserve un **état interne** et un **historique** (section 1.2.1.2), ce qui lui confère une **mémoire contextuelle** et améliore la **cohérence** de l'apprentissage sur le long terme.

Écosystèmes et Coopérations Symbiotiques

Les **écosystèmes** fournissent de nombreux exemples de **coopération** interspécifique. Les insectes pollinisent les plantes, les lichens résultent de la symbiose entre algues et champignons, et la mycorhize associe champignons et racines de plantes. Dans chacun de ces cas, les organismes obtiennent un **bénéfice mutuel**, qu'il s'agisse d'un meilleur accès aux ressources, d'une protection renforcée ou d'une plus grande capacité d'adaptation. Cette dynamique repose sur la **notion de "gain commun"**, qui illustre la **synergie**. La coexistence de deux entités ou espèces génère une **valeur ajoutée** absente lorsqu'elles agissent isolément.

Dans le **Deep Synergy Learning (DSL)**, ce principe s'incarne à travers la **mesure de synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$ (voir section 1.2.2) et l'**ajustement** adapté des pondérations $\omega_{i,j}(t)$. De la même façon que deux espèces coopérantes se renforcent l'une l'autre, deux entités informationnelles voient leurs liens se consolider lorsqu'elles interagissent efficacement. Les **écosystèmes** diversifiés, riches en symbioses, font preuve d'une **résilience** considérable face aux menaces comme la sécheresse ou la prédation, grâce aux ressources complémentaires que chaque espèce apporte. Un **réseau synergique (DSL)** regroupant des entités variées, qu'elles soient visuelles, textuelles ou auditives, développe une plus grande **robustesse** et une meilleure **flexibilité**. En modulant continuellement ses interactions, il s'adapte plus efficacement aux imprévus et aux évolutions de son environnement.

Synchronisation Collective : Bancs de Poissons et Nuées d'Étourneaux

Les **bancs de poissons** et les **vols d'oiseaux**, comme les nuées d'étourneaux, illustrent un phénomène remarquable de **synchronisation**. De larges groupes se déplacent de manière presque chorégraphiée, sans chef unique, chaque individu ajustant sa trajectoire en fonction de celle de ses voisins. Ce processus engendre un **effet émergent** de cohésion et d'harmonisation.

Ce type d'organisation repose généralement sur quelques **règles simples**, comme l'alignement des vitesses, le maintien d'une distance de sécurité et l'attraction mutuelle. Une fois combinées, ces

règles produisent des comportements collectifs complexes. Dans le cadre du **Deep Synergy Learning (DSL)**, la **synergie** et la **mise à jour** des liens jouent un rôle similaire. Lorsque la coopération entre deux entités s'avère bénéfique, elles se **synchronisent**, ce qui contribue à optimiser le **réseau global**.

Les bancs ou les nuées font également preuve d'une grande **plasticité**, se reconfigurant rapidement face à un prédateur ou un obstacle. De façon parallèle, un **réseau synergique** peut, à tout instant, **adapter** sa structure dès lors que le contexte ou les données évoluent, sans nécessiter de "réentraînement" global et figé.

Conclusion

Les exemples précédents – colonies d'insectes, cerveau, écosystèmes ou synchronisations collectives – démontrent que la synergie émerge lorsque des entités locales et relativement simples **coopèrent** selon des **règles d'interaction** et d'**adaptation**. Sans supervision centrale et sans plan préconçu, il se forme souvent des **structures** ou des **comportements** remarquablement organisés et robustes, capables de s'ajuster aux contraintes du milieu.

Ces observations, empruntées à la nature, guideront la formulation plus formelle des **algorithmes** et la mise en place de **protocoles d'évaluation** pour le DSL. Dans les chapitres suivants, nous verrons comment concrétiser ces analogies sous la forme de modèles mathématiques, de règles de mise à jour et d'applications pratiques, visant à faire du Deep Synergy Learning un **paradigme opérationnel** pour une **IA forte** (ou du moins plus autonome et plus générale).

1.3. Importance de l'Auto-Organisation

Dans le cadre du **Deep Synergy Learning (DSL)**, l'un des aspects les plus novateurs réside dans sa capacité d'**auto-organisation**. Alors que les approches d'IA traditionnelles s'appuient souvent sur des architectures hiérarchiques fixes, le DSL encourage un processus dynamique où les **entités d'information** établissent et révisent en continu leurs relations, en fonction de la **synergie** qu'elles détectent entre elles (voir sections précédentes). Cette section (1.3) aborde l'importance de l'auto-organisation sous plusieurs angles, depuis les sources d'inspiration biologiques (1.3.1) jusqu'à l'impact sur la résilience des modèles (1.3.6) et les perspectives globales (1.3.7).

1.3.1. Inspirations Biologiques et Cognitives

L'auto-organisation, au sens où l'entend le DSL, n'est pas un concept nouveau. En réalité, de nombreux **systèmes naturels** (organismes vivants, écosystèmes, cerveau, etc.) démontrent d'extraordinaires facultés d'organisation **spontanée** et **adaptative**. Le **Deep Synergy Learning** cherche à exploiter ces principes, en transposant aux systèmes informatiques des mécanismes qui ont fait leurs preuves dans le vivant.

1.3.1.1. Le cerveau comme paradigme de plasticité



Le cerveau humain compte des centaines de milliards de neurones, reliés par des synapses dont les forces (poids synaptiques) évoluent selon l'activité, un phénomène désigné par “plasticité synaptique”. On peut symboliser ce principe via une équation de type Hebbien généralisée, par exemple :

$$w_{ij}(t+1) = w_{ij}(t) + \eta[A_{ij}(t) - \lambda w_{ij}(t)],$$

où w_{ij} est la “force” de la synapse reliant le neurone i au neurone j , η un taux d'apprentissage, $A_{ij}(t)$ l'activité conjointe des neurones i et j (ou une mesure de leur association), et λ un terme de régulation. L'important est de voir comment la synapse se renforce si l'activité commune est

significative, rappelant la **mise à jour** des **pondérations synergiques** dans le DSL (cf. section 1.2.6).

Dans le cerveau, des regroupements (ou *assemblées*) de neurones se forment lorsque leurs synapses se renforcent mutuellement. Le concept de cluster auto-organisé dans le DSL (où les entités se regroupent pour exploiter leur synergie) ressemble à cette idée :

$$\mathcal{C}^* = \operatorname{argmax}_{\mathcal{C} \subset \{\mathcal{E}_1, \dots, \mathcal{E}_n\}} \sum_{(i,j) \in \mathcal{C} \times \mathcal{C}} \omega_{i,j},$$

c'est-à-dire la recherche d'un sous-ensemble \mathcal{C} maximisant la somme des pondérations entre ses membres, analogiquement à des neurones qui s'assemblent pour traiter un même stimulus.

La cognition résulte du fait que certains assemblages persistent dans le temps et renforcent leur "*autoroute synaptique*" (par exemple, mémorisation d'un stimulus répété). Dans le DSL, l'émergence de **micro-réseaux** stables autour d'entités synergiques reflète ce même phénomène d'**apprentissage distribué**, favorisant la récurrence de schémas pertinents.

1.3.1.2. Colonies d'insectes et intelligence collective

Les colonies de fourmis s'illustrent souvent par leur capacité à trouver des chemins optimaux pour la recherche de nourriture. Ce comportement repose sur des **interactions locales** (dépôt de phéromones, etc.), et non sur un plan global. L'analogie avec le DSL se fait sentir lorsque l'on considère que chaque fourmi incarne une "entité" prenant des décisions locales qui, agrégées, produisent une optimisation globale.



On peut modéliser la concentration de phéromones le long d'un chemin par une fonction $\phi(t)$ obéissant à un mécanisme de renforcement/dissipation, par exemple :

$$\frac{d\phi}{dt} = \alpha I(\text{fourmis}) - \beta \phi,$$

où α modélise l'ajout de phéromones (si des fourmis passent) et β la vitesse d'évaporation. Le DSL retrouve un concept similaire quand il gère l'**intensité** d'une liaison synergique via des règles de renforcement ou de décroissance (cf. équations de mise à jour des pondérations, section 1.2.4).

À l'instar de chemins de phéromones qui “naissent” (s'ils sont utiles) et “disparaissent” (s'ils ne sont plus empruntés), les connexions synergiques $\omega_{ij}(t)$ du DSL se renforcent ou s'étiolent, entraînant la **création** ou la **rupture** de clusters d'entités d'information. Cette dynamique **auto-organisée** ne requiert pas de commande centrale.

1.3.1.3. Systèmes dynamiques et attracteurs auto-organisés

Exemples de modèles dynamiques

De nombreux systèmes auto-organisés (réacteurs chimiques de type Belousov-Zhabotinsky, réactions oscillantes, etc.) peuvent s'exprimer par des **équations différentielles** à plusieurs variables, où les interactions locales créent des **boucles de rétroaction** positives ou négatives. Les motifs qui émergent (ondes chimiques, spirales...) sont des exemples d'**attracteurs** dynamiques.

Analogies avec le graphe synergique

Dans le DSL, on peut considérer que l'**état global** du réseau est donné par la matrice $\Omega(t) = [\omega_{i,j}(t)]$. L'évolution de $\Omega(t)$ suit des règles de type :

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)].$$

Cette équation peut se voir comme un **système dynamique** discret, susceptible de converger vers un attracteur (une configuration stable) ou d'entrer dans des régimes oscillatoires, reflétant l'**auto-organisation**.

Stabilité et oscillations

Dans un cadre mathématique plus poussé, on peut étudier la **stabilité** de l'état Ω^* en résolvant :

$$\omega_{i,j}(t+1) - \omega_{i,j}(t) = 0,$$

pour tous (i, j) . On obtient alors une condition d'équilibre, ce qui, dans un contexte du DSL, correspond à un réseau où chaque liaison a atteint une valeur stable en regard des synergies (sous réserve que l'on n'introduise pas de nouvelles entités ou de nouvelles données). Une **analyse de la Jacobienne** locale pourrait prédire si cet état Ω^* est un **point d'attraction**, un **cycle limite** ou un **chaos**.

1.3.1.4. Vers un transfert de ces principes au DSL

Les systèmes biologiques et cognitifs décrits ci-dessus mettent en avant quatre qualités fondamentales :

- **Plasticité** (adaptation continue)

- **Robustesse** (résistance au bruit ou à la perturbation)
- **Emergence** (apparition de structures non planifiées)
- **Auto-régulation** (équilibre entre renforcement et inhibition)

Le DSL s’inspire de ces qualités pour proposer un **apprentissage distribué**, dans lequel les entités d’information se comportent comme des agents qui renforcent leurs liens lorsqu’ils détectent un **gain** (cf. notion de synergie) et les relâchent sinon. Matériellement, on transpose ces mécanismes via :

$$\omega_{i,j}(t+1) = \omega_{i,j}(t)\eta\left(S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)\right),$$

et des **clusters** se forment spontanément si le gain est significatif.

Contrairement aux réseaux neuronaux profonds classiques, où seule la valeur des poids évolue dans une architecture figée, l’**auto-organisation** du DSL laisse la structure du réseau se **recomposer** en fonction des synergies détectées. Les entités elles-mêmes peuvent évoluer, en modifiant leur **représentation** ou leur **état** interne, et donc modifier indirectement la dynamique globale du réseau.

1.3.1.5. Conclusion partielle : fondements biologiques de l’auto-organisation

Cette section a illustré, par des références à la **plasticité synaptique**, à l’**intelligence collective** chez les insectes et à divers **systèmes dynamiques**, l’importance des mécanismes d’auto-organisation en biologie et en cognition. Le **Deep Synergy Learning** cherche à exploiter ces mêmes principes pour créer des modèles d’intelligence artificielle moins dépendants d’une structure hiérarchique imposée, plus adaptatifs, et potentiellement plus proches d’une **IA générale**. Les sections suivantes (1.3.2 à 1.3.7) mettront en évidence des concepts plus généraux d’**émergence**, de **feedback**, et détailleront comment l’auto-organisation du DSL se compare aux méthodes d’apprentissage classiques, avant de conclure sur la pertinence de cette approche pour des architectures globales et multimodales.

1.3.2. Concepts Clés : Émergence, Auto-Régulation, Feedback

La dynamique **auto-organisée** que promeut le Deep Synergy Learning (DSL) repose sur plusieurs notions fondamentales, à la fois conceptuelles et mathématiques, qui permettent de comprendre comment un système d’entités peut s’ajuster et se coordonner sans planification centrale. Trois d’entre elles sont particulièrement essentielles : l’**émergence**, l’**auto-régulation** et le **feedback** (rétroaction). Les sous-sections ci-après détaillent la signification de chacun de ces concepts et leur importance dans le cadre du DSL.

1.3.2.1. Émergence

Définition. Le terme *émergence* décrit la situation où un **phénomène global** ou une **structure** à l'échelle macroscopique naît de l'interaction de **composants** à l'échelle microscopique, sans qu'il n'y ait de plan explicite ni de chef d'orchestre. Par exemple, la forme d'une nuée d'oiseaux ou l'organisation complexe d'une ruche peuvent être vus comme des propriétés émergentes d'un grand nombre d'interactions locales.

Dans le Deep Synergy Learning, l'**émergence** se manifeste par la formation de **clusters** (ou micro-réseaux) d'entités d'information, que l'on peut qualifier de “sous-systèmes cohérents” au sein du grand réseau synergique. Ces **clusters** ne sont pas prédéfinis. Ils **apparaissent** et parfois **disparaissent** en fonction de la synergie détectée entre les entités.

Un cluster $\mathcal{C} \subset \{\mathcal{E}_1, \dots, \mathcal{E}_n\}$ peut s'illustrer par une somme de pondérations internes $\sum_{(i,j) \in \mathcal{C} \times \mathcal{C}} \omega_{i,j}$ qui se trouve **maximalisée** à un certain moment t .

De plus, ces clusters émergents peuvent générer de nouvelles **représentations** ou entités “conjointes”, lesquelles elles-mêmes nourrissent l'**évolution** du réseau.

Sur le plan formel, on peut lier l'émergence au fait que la **dynamique** des pondérations synergiques $\omega_{i,j}(t)$ conduit le système vers des **attracteurs** (états stables ou semi-stables). L'apparition d'un cluster correspond alors à un **bassin d'attraction** particulier, où les liens internes se renforcent suffisamment pour maintenir l'entité groupée, un phénomène typique des **systèmes dynamiques complexes**.

1.3.2.2. Auto-Régulation

Dans un système auto-organisé, chaque composant ou lien s'**ajuste** en fonction de **règles locales**. Lorsque la coopération est bénéfique, le lien se renforce, tandis que s'il est jugé néfaste ou inutile, il s'affaiblit. Ce **mécanisme d'ajustement** relève de l'**auto-régulation**, car il ne dépend pas d'une autorité externe qui dirait “ces entités doivent se connecter, celles-ci doivent se séparer”.

Comme décrit en section 1.2.4, une règle fréquemment utilisée dans le DSL pour modéliser l'évolution des pondérations synergiques $\omega_{i,j}(t)$ est :

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)].$$

- η représente un taux d'apprentissage (contrôle la rapidité de l'ajustement).
- $S(\mathcal{E}_i, \mathcal{E}_j)$ est la **synergie** calculée entre les entités i et j .
- τ agit comme un terme de décroissance ou de “frottement”, évitant que les pondérations ne croissent indéfiniment.

Le paramètre τ et le choix de la fonction $S(\cdot)$ permettent d'**éviter les dérives** (ex. explosion des connexions) et de stabiliser le système dans un régime d'**énergie** ou de **coût** relativement bas (voir fonction \mathcal{J} discutée en 1.2.6). Ce mécanisme d'**auto-régulation** dispense de toute redéfinition

manuelle de la topologie du réseau. Il s'adapte de lui-même aux contraintes et aux données entrantes.

1.3.2.3. Feedback (Rétroaction)

Le **feedback** désigne le processus par lequel la **sortie** d'un système ou l'**état** d'une partie du système **retourne** sur lui-même, influençant à son tour la configuration ou le comportement du système. Dans le **DSL**, le feedback se manifeste principalement à travers l'**apprentissage** des synergies. La manière dont deux entités ont coopéré à l'instant t influence la pondération $\omega_{i,j}(t+1)$, qui déterminera ensuite leur coopération future.

Le **feedback positif** est un mécanisme par lequel une forte synergie entre deux entités, notée $S(\mathcal{E}_i, \mathcal{E}_j)$, conduit à un renforcement mutuel de leur connexion. En d'autres termes, si la synergie initiale entre deux entités est élevée, leur interaction devient progressivement plus forte, renforçant encore davantage leur synergie. Ce processus peut continuer jusqu'à atteindre un certain plafond, au-delà duquel la connexion ne peut plus croître, stabilisant ainsi leur relation synergique.

À l'inverse, le **feedback négatif** intervient lorsque la synergie entre deux entités est faible ou nulle, voire négative. Dans ce cas, la pondération de leur lien décroît progressivement, réduisant la probabilité qu'elles interagissent à l'avenir. Cependant, ce mécanisme n'est pas définitif. Un changement de contexte ou d'état peut, à terme, améliorer la compatibilité entre deux entités et rétablir leur capacité à collaborer ou à former une connexion.

Outre les liens binaires (entre paires d'entités), il peut exister des **boucles de rétroaction plus complexes** à l'échelle de plusieurs clusters. Par exemple, un cluster \mathcal{C}_1 peut influencer la formation ou la dissolution d'un autre cluster \mathcal{C}_2 en modifiant le contexte, les ressources partagées, etc. Cela équivaut à considérer des **fonctions de synergie n-aire**, où les "coûts" ou "gains" de chaque liaison dépendent aussi des états des autres liens et entités.

$$S(\mathcal{E}_i, \mathcal{E}_j \mid \{\mathcal{E}_k\}_{k \neq i,j}), \quad \omega_{i,j}(t+1) = f(\omega_{i,j}(t), \{S(\mathcal{E}_i, \mathcal{E}_j \mid \dots)\}, \dots).$$

1.3.2.4. Synthèse : de la dynamique locale à la cohérence globale

Les trois concepts d'**émergence**, d'**auto-régulation** et de **feedback** interagissent étroitement dans le Deep Synergy Learning :

- L'**émergence** se produit parce que les liens (pondérations synergiques) se développent localement sans qu'aucun schéma global ne soit imposé.
- L'**auto-régulation** assure que chaque lien est entretenu ou affaibli en fonction de la **pertinence** mesurée (synergie), maintenant ainsi une forme de **plasticité** dans l'ensemble du réseau.
- Le **feedback** perpétuel, tant au niveau des paires d'entités que des ensembles plus larges, crée un **cycle** où l'état du réseau à l'instant t influence les coopérations à l'instant $t+1$, et ainsi de suite.

Ce mode de fonctionnement "bouclé" et distribué est très différent de la logique hiérarchique classique (voir 1.2.4). Il rappelle davantage les **systèmes biologiques** et autres **systèmes**

complexes, où la stabilité et les structures apparaissent comme le fruit d'interactions locales répétées, produisant parfois des schémas surprenants d'ordre (ou de désordre) global.

1.3.2.5. Conclusion intermédiaire : vers une IA plus adaptative

Comprendre les notions d'**émergence**, d'**auto-régulation** et de **feedback** est crucial pour saisir pourquoi le **DSL** peut offrir des capacités d'adaptation et de créativité supérieures à celles des architectures figées. Loin d'être un simple attribut de "design" théorique, ces concepts constituent la **dynamique interne** qui permet au réseau synergique de se **reconfigurer** en permanence, de **fusionner** ou de **séparer** des entités en clusters, et d'**apprendre** sans devoir figer sa topologie.

Dans la section suivante (1.3.3), nous mettrons en perspective ces mécanismes d'auto-organisation avec les **méthodes d'apprentissage classiques**, afin d'explorer les gains potentiels (et les défis) qu'apporte l'introduction de ces idées dans le champ de l'intelligence artificielle.

1.3.3. Comparaison avec les Méthodes d'Apprentissage Classiques

Les principes d'**auto-organisation** et de **synergie** du Deep Synergy Learning (DSL) contrastent fortement avec les schémas employés par les méthodes d'apprentissage les plus courantes, qu'il s'agisse d'algorithmes supervisés, non supervisés ou même de certaines approches en renforcement. Dans la présente section, nous mettons en lumière les principales différences entre ces paradigmes "classiques" (tels que le **Deep Learning** hiérarchique, la **clustering** statique ou la **rétropropagation de gradient** end-to-end) et le **DSL**, qui s'appuie avant tout sur une **construction dynamique** du réseau via la synergie informationnelle.

1.3.3.1. Apprentissage supervisé traditionnel

Dans un **cadre supervisé** classique, on dispose d'un ensemble de données $\{(\mathbf{x}_i, y_i)\}_{i=1, \dots, N}$, où \mathbf{x}_i constitue l'entrée (vecteur de caractéristiques, image, etc.) et y_i le label associé (catégorie, valeur numérique, etc.). Un **modèle** (par exemple, un réseau de neurones) est alors **entraîné** à minimiser une **fonction de coût** :

$$\mathcal{L}(\theta) = \frac{1}{N} \sum_{i=1}^N L(f_{\theta}(\mathbf{x}_i), y_i),$$

où f_{θ} désigne le modèle paramétré par θ (poids et biais), et $L(\cdot)$ une fonction de perte (entropie croisée, MSE, etc.).

La plupart du temps, l'**architecture** du réseau est **figée** : on choisit un nombre de couches, une topologie (CNN, MLP, etc.) et on la conserve durant tout l'entraînement. La **rétropropagation** permet d'ajuster les poids internes θ , sans pour autant réorganiser la **structure** (aucune suppression ni création de neurones, aucune reconfiguration des connexions).

Comparaison avec le DSL.

Dans un réseau **DSL**, on ne se limite pas à la mise à jour des paramètres : on autorise la **topologie** à évoluer (création, renforcement, fusion, ou rupture de liens) en se fondant sur la **mesure de synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$. Dans un modèle supervisé classique, la structure reste **inchangée** tout au long de l'apprentissage.

Dans l'approche **supervisée** traditionnelle, l'effort se concentre sur la **réduction de l'erreur** par rapport aux labels y . À l'inverse, le **DSL** se concentre sur la **relation** entre les différentes données elles-mêmes. La **formation** de clusters et l'**auto-organisation** peuvent émerger même en l'absence de labels, dès lors que le réseau détecte des synergies fructueuses.

Un **modèle supervisé** doit souvent être **réentraîné** (ou au moins fine-tuné) dès lors qu'apparaissent de nouveaux types de données ou qu'une **distribution** se modifie. Dans le DSL, la structure s'**auto-réorganise** en intégrant ou en écartant certaines entités, en fonction de la synergie relevée, ce qui apporte davantage de **souplesse** et de **plasticité** face aux environnements changeants.

1.3.3.2. Apprentissage non supervisé et clustering statique

Dans l'apprentissage **non supervisé**, on cherche généralement à regrouper des points \mathbf{x}_i en clusters. Des méthodes comme **k-means** démarrent avec un nombre de clusters (k) et une mesure de distance (par exemple euclidienne ou cosinus), puis procèdent à une série d'itérations visant à minimiser

$$\sum_{i=1}^N \min_{1 \leq c \leq k} \|\mathbf{x}_i - \boldsymbol{\mu}_c\|^2,$$

où $\boldsymbol{\mu}_c$ sont les centres de clusters. Dans **k-means**, le paramètre k doit être défini **a priori**, et même si d'autres algorithmes (tels que DBSCAN ou le clustering agglomératif) peuvent estimer le nombre de groupes, ils reposent néanmoins sur une **logique statique**, c'est-à-dire qu'on exécute l'algorithme une fois et on obtient une unique partition.

Dans ces approches classiques, les points ou vecteurs ne possèdent pas de **dynamique** interne. Leur représentation reste figée, et les liens potentiels entre eux n'évoluent pas au fil du temps en fonction d'une éventuelle collaboration. Le **DSL** se distingue précisément sur ce point. Les **clusters** peuvent **apparaître** ou **se dissoudre** de manière **progressive** selon les pondérations synergiques $\omega_{i,j}(t)$, sans qu'un paramètre k fixe ne contraigne le regroupement. La structure du réseau se **façonne** ainsi en continu, guidée par la dynamique d'apprentissage.

En outre, chaque entité \mathcal{E}_i du réseau DSL peut être **active**, c'est-à-dire posséder un **état interne** et évoluer au cours du temps. Dans un clustering **statique**, les points demeurent figés dans leurs coordonnées, sans aucun mécanisme de feedback local ou de synergie émergente. Enfin, lorsqu'apparaissent de nouvelles données ou qu'un changement de distribution survient, le DSL est capable de **réévaluer** les synergies et de **reconfigurer** automatiquement les clusters, tandis qu'un clustering non supervisé traditionnel exigerait la relance complète de l'algorithme, ignorant dans la foulée toute information antérieure.

1.3.3.3. Apprentissage par renforcement (RL)

Dans le **Reinforcement Learning**, un **agent** interagit avec un environnement, perçoit des **récompenses** (positives ou négatives) selon ses actions, et cherche à **maximiser** un **cumul de récompenses**. Les algorithmes de RL (Q-learning, SARSA, méthodes basées sur une politique) visent souvent à estimer la fonction de valeur $Q(s, a)$ ou la politique optimale $\pi(a|s)$. De nombreux scénarios de RL (par ex. Deep Q-Network) s'appuient sur un **réseau profond** paramétré, dont les poids sont ajustés via un **objectif** portant sur la récompense cumulée.

Dans la plupart des implémentations **RL** traditionnelles, la **structure** du réseau reste figée. Au cours de l'apprentissage, ni le nombre de couches ni la forme du modèle ne sont modifiés.

Synergie vs. récompense.

Alors que le RL privilégie l'action menant à la **maximisation** de la récompense, le DSL promeut la **coopération** (ou la connexion) qui maximise la **synergie**. Il est envisageable de développer un **DSL-RL** où la **synergie** intervient comme une composante de la **fonction de récompense**. Plus la **collaboration** entre entités s'avère fructueuse, plus les liens correspondants se renforcent.

Reconfiguration continue.

Dans le RL classique, on n'ajoute ni ne supprime spontanément des neurones ou des liaisons durant l'apprentissage. À l'inverse, le DSL autorise une **recomposition** de l'architecture si cela contribue à améliorer la performance (ou la cohérence). Les pondérations peuvent émerger, se renforcer, ou disparaître au fil du temps.

Apprentissage distribué.

On peut voir les entités du DSL comme autant d'"agents partiels" qui s'accordent via la **synergie**, au lieu de recourir à un **agent unique** muni d'un policy network fixe. Cette perspective se distingue du **RL centralisé**. L'organisation émerge de l'**interaction locale** entre entités au lieu de suivre une stratégie unique imposée à l'ensemble du réseau.

1.3.3.4. Synthèse des contrastes fondamentaux

Les approches classiques, qu'elles soient **supervisées**, **non supervisées** ou basées sur le **renforcement**, reposent généralement sur une **architecture** ou un **ensemble d'hypothèses** prédéfinies. Cela peut correspondre à un réseau feed-forward, à un nombre fixé de clusters à détecter ou à une paramétrisation spécifique. À l'inverse, le **DSL** ne fixe pas la structure à l'avance ; il permet à celle-ci de **se former** et de **se modifier** de manière autonome, simplement en s'appuyant sur les mesures de **synergie** et sur les **règles** d'auto-organisation.

Dans les méthodes classiques, on privilégie soit le **rapprochement** entre données et labels (apprentissage supervisé), soit la **répartition** de points dans des groupes (apprentissage non supervisé). Le DSL, au contraire, met l'accent sur la **relation** entre les entités, qu'il s'agisse de leur **coexistence** ou de leur **synergie**, et c'est ce mécanisme relationnel qui détermine l'organisation d'ensemble.

Ce renversement de perspective — considérer la donnée non pas comme un **input passif**, mais comme un **acteur** recherchant activement des **connexions** — constitue la marque distinctive du DSL. Les algorithmes traditionnels sont souvent conçus pour un entraînement **en batch**. Les données sont d’abord collectées, le modèle est ensuite entraîné, puis sa configuration est figée. Des versions en ligne existent, mais elles n’ont pas pour vocation de retoucher la **structure** du modèle ; elles n’agissent que sur les poids.

Le **DSL**, en revanche, s’adapte naturellement à un **flux continu** de données. À mesure que de nouvelles informations apparaissent, le réseau peut se **reconstruire** pour s’ajuster aux conditions changeantes ou intégrer de **nouvelles entités**. Il favorise ainsi une **adaptation permanente**, plutôt que de se conformer à un schéma fixe déterminé à l’avance.

1.3.3.5. Conclusion : l’auto-organisation comme rupture de paradigme

La comparaison avec les méthodes classiques met en évidence le caractère **révolutionnaire** de l’auto-organisation prônée par le **DSL**. La capacité à **faire évoluer la carte du réseau** plutôt que de se limiter à des architectures fixes, l’intégration d’une **synergie** mesurée entre entités, la facilité d’**apprentissage continu** et la résistance aux changements de distribution en font une approche fondamentalement différente.

Ces différences ne signifient pas que le DSL vienne nécessairement **remplacer** l’ensemble des paradigmes existants. On peut au contraire envisager des **approches hybrides** qui combinent la puissance des algorithmes classiques (supervisés ou RL) avec la **plasticité** et la **coévolution** propres au DSL. Les sections qui suivront (1.3.4 à 1.3.7) approfondiront l’idée d’une **auto-organisation multimodale**, ainsi que l’**impact** concret de cette approche sur la robustesse et la résilience des modèles, pour finalement ouvrir la voie à une IA plus générale et plus souple.

1.3.4. Le Rôle des Flux d’Information Multimodaux

Parmi les avantages majeurs de l’**auto-organisation** telle que proposée dans le **Deep Synergy Learning (DSL)** figure la capacité à gérer efficacement des **flux d’information** variés (image, texte, audio, données sensorielles, etc.). En effet, l’idée de **synergie informationnelle** (développée aux sections 1.2.2 et 1.2.3) s’applique de manière particulièrement féconde dès lors que plusieurs modalités sont en jeu, car chaque flux peut **renforcer** ou **compléter** les autres. Cette sous-section (1.3.4) illustre la manière dont l’auto-organisation des entités d’information s’adapte à la **multi-modalité**, et pourquoi cela peut conférer au DSL une **puissance d’intégration** hors de portée des architectures classiques, souvent cloisonnées.

1.3.4.1. Notion de multi-modalité dans le DSL

Dans un système multimodal, on distingue plusieurs sources de données :

- **Modalité visuelle** : images, vidéos (entité $\mathcal{E}_{\text{visuelle}}$)

- **Modalité auditive** : signaux sonores, musique, parole (entité $\mathcal{E}_{\text{auditive}}$)
- **Modalité textuelle** : séquences de mots, documents, balises sémantiques (entité $\mathcal{E}_{\text{textuelle}}$)
- **Modalités sensorielles complémentaires** : capteurs physiques (température, distance, pression), signaux biométriques, etc.

Dans le **Deep Synergy Learning**, chaque flux peut être représenté par une ou plusieurs **entités d'information** (voir section 1.2.1). Ainsi, un flux visuel pourra donner naissance à plusieurs entités – par exemple, une entité pour la “carte de caractéristiques” (feature map) d’une image, une autre pour son histogramme de couleurs, etc. Ces entités ne sont pas limitées à une seule couche. Elles peuvent interagir avec toutes les autres entités, qu’elles soient visuelles, textuelles ou auditives, en fonction de leur **synergie**.

1.3.4.2. Avantages de la synergie entre modalités

La **synergie** est particulièrement critique en multi-modalité :

$$S(\mathcal{E}_{\text{visuelle}}, \mathcal{E}_{\text{auditive}}) > 0 \\ \Rightarrow \text{l'information visuelle et l'information auditive se complètent efficacement.}$$

Par exemple, dans une scène vidéo où l’on entend quelqu’un parler, la composante auditive (voix) et la composante visuelle (mouvements de lèvres, expressions faciales) coopèrent pour renforcer la compréhension globale.

Lorsqu’une modalité est dégradée (par exemple, l’audio est parasité par du bruit de fond), une autre modalité (l’image) peut prendre le relais, augmentant la **tolérance aux perturbations**. Dans un **réseau synergique**, cela se traduit par une réduction dynamique du poids $\omega_{\text{audio,visuel}}(t)$ si l’audio devient momentanément peu fiable, et par le renforcement d’autres liens synergiques.

Les entités associées à différentes modalités peuvent créer des **clusters** communs si leur synergie mutuelle est élevée. Mathématiquement, on peut repérer un cluster \mathcal{C}_m formé d’entités $\{\mathcal{E}_{\text{visuelle}}^a, \mathcal{E}_{\text{auditive}}^b, \mathcal{E}_{\text{textuelle}}^c\}$ tel que

$$\sum_{(i,j) \in \mathcal{C}_m \times \mathcal{C}_m} \omega_{i,j} > \theta_{\text{seuil}},$$

où θ_{seuil} est un paramètre symbolisant la cohésion requise pour se constituer en cluster. Ce regroupement favorise l’**intégration sémantique** de différents types de signaux.

1.3.4.3. Mécanismes d’auto-organisation pour la fusion multimodale

Dans un réseau hiérarchique classique, la fusion multimodale est généralement réalisée dans une “couche” dédiée (p. ex. concaténation de features). Dans le DSL, la fusion naît **spontanément** :

$$\omega_{\text{visuelle,textuelle}}(t+1) = \omega_{\text{visuelle,textuelle}}(t) \eta [S(\mathcal{E}_{\text{visuelle}}, \mathcal{E}_{\text{textuelle}}) - \tau \omega_{\text{visuelle,textuelle}}(t)].$$

Plus la coopération entre les deux modalités est jugée bénéfique (ex. augmentation de la performance ou de l'information mutuelle), plus ce lien se consolide.

Le caractère **évolutif** d'un **cluster multimodal** est essentiel. Selon le **contexte**, la modalité audio peut jouer un rôle plus ou moins important, comme dans le cas d'un concert de musique par rapport à un environnement silencieux. Les liens synergiques inter-modaux varient au fil du temps, entraînant l'émergence ou la dissolution de sous-ensembles cohérents.

On peut également enrichir la mesure de synergie S par un **contexte** $c(t)$, de sorte qu'il y ait des synergies conditionnelles :

$$S(\mathcal{E}_i, \mathcal{E}_j \mid c(t)).$$

Dans une scène nocturne, par exemple, la vision est moins fiable, tandis que l'audio reste pertinent. Le contexte "nuit" pourrait diminuer certaines liaisons visuelles et renforcer des liaisons audio ou thermiques, mettant l'accent sur la modalité la plus fiable.

1.3.4.4. Exemples concrets et gains opérationnels

Imaginez un système de surveillance qui reçoit en continu des flux vidéo, audio, et des capteurs de mouvement. Dans une approche classique, il faudrait concevoir à l'avance un schéma de fusion ; dans le DSL, les **entités** correspondant à chaque flux s'**auto-organisent** pour détecter des patrons communs (ex. détection d'un intrus combinant une silhouette anormale et des sons inhabituels).

Le **renforcement** des liens entre un flux "caméra infrarouge" et un flux "microphones" peut se produire si ces deux modalités se complètent pour la détection nocturne.

Dans le cadre d'interfaces multimodales (voix, gestes, expressions faciales), chaque modalité peut être captée par une entité DSL. Si l'utilisateur parle tout en faisant des gestes, le cluster "voix + gestes" obtient un fort score de synergie, permettant de mieux interpréter la signification globale (commande ou intention).

L'**auto-organisation** permet au système de **s'adapter** aux préférences de l'utilisateur. Si sa diction est peu claire mais que ses gestes sont très expressifs, les pondérations vers la modalité gestuelle se renforceront spontanément.

Le **DSL** peut traiter des couples **image + légende textuelle**. Une entité **extraite de texte** peut renforcer sa synergie avec une entité **détecteur d'objets** dans l'image, en particulier lorsque certains mots clés coïncident avec des formes reconnues.

Les entités "texte" et "image" peuvent alors former un **cluster** spécifique de type "concept visuel + label textuel", facilitant les tâches d'annotation automatique ou de recherche d'images par mots-clés (et vice versa).

1.3.4.5. Forces et défis de l'approche synergique multimodale

L'**absence** de couche de fusion prédéfinie confère une grande **flexibilité**. Plutôt que d'imposer un schéma fixe de combinaison entre les flux visuel, audio ou textuel, le **DSL décide** de manière autonome quelles modalités sont les plus pertinentes à associer. Cette adaptabilité renforce également la **robustesse**. Si un flux devient bruité ou indisponible, comme un microphone défectueux, le système peut se reconfigurer et s'appuyer davantage sur d'autres canaux. Enfin, l'**auto-organisation** permet la **découverte** de combinaisons inédites. Des modalités supposées peu corrélées peuvent révéler une forte synergie dans un certain contexte, ce qu'un pipeline classique plus rigide aurait pu ignorer.

En contrepartie, la **complexité computationnelle** peut croître rapidement à mesure que le nombre de modalités et d'entités augmente. L'évaluation du **gain de synergie** entre plusieurs canaux peut s'avérer coûteuse, rendant nécessaire l'utilisation d'**heuristiques** ou d'**approximations** pour maintenir une efficacité raisonnable. De plus, l'enchaînement rapide de contextes (passer d'une scène de rue à une scène d'intérieur, par exemple) peut induire une **volatilité** élevée des connexions synergiques ; il faut alors recourir à des **coefficients de régulation** (τ) ou à des dispositifs de mémorisation pour **stabiliser** la configuration. Enfin, l'**interopérabilité** entre flux implique de disposer de **représentations** comparables ou d'un espace commun adapté (par exemple, une même métrique pour estimer la similarité entre un vecteur d'image et un vecteur de texte).

1.3.4.6. Conclusion : la multi-modalité comme terrain idéal pour le DSL

Le **Deep Synergy Learning** apparaît particulièrement adapté à l'intégration de **flux multimodaux**. Là où les méthodes classiques imposent généralement une architecture de fusion fixe (par exemple, concaténer l'output d'un CNN d'image et d'un RNN de texte), le DSL laisse les **entités** (issues de différentes modalités) **explorer** leurs synergies possibles. Les liens synergiques se **cristallisent** ou se **dissolvent** en fonction de leur valeur ajoutée, permettant une **adaptation dynamique** aux contextes et aux sources de bruit.

Cette approche auto-organisée favorise l'**émergence** (voir section 1.3.2) de clusters multimodaux plus ou moins stables, qui peuvent évoluer au fil du temps. Sur le plan applicatif, il en résulte une **robustesse** et une **flexibilité** uniques, aussi bien pour la reconnaissance d'événements complexes, la détection d'anomalies, ou encore la mise en place d'interactions homme-machine riches. Les sections suivantes (1.3.5, 1.3.6, 1.3.7) examineront l'**évolution dynamique** de ces réseaux, l'**impact** sur la **résilience** des modèles, et la perspective d'une **approche plus globale** de l'apprentissage.

1.3.5. Évolution Dynamique et Adaptation en Continu

Le **Deep Synergy Learning (DSL)** se distingue des approches classiques de l'IA par sa capacité à laisser les entités d'information et leurs connexions **évoluer** en permanence, au gré de la synergie détectée entre les flux de données. Cette caractéristique d'**évolution dynamique** et d'**adaptation en continu** s'inspire directement de divers phénomènes naturels (cerveau, écosystèmes, colonies d'insectes...) et confère au DSL une **plasticité** inhabituelle dans le domaine de l'apprentissage

automatique. Dans cette section, nous analysons pourquoi cette évolution constante est essentielle, quels sont ses principes fondamentaux et comment elle se formalise mathématiquement.

1.3.5.1. Principes généraux de l'adaptation continue

Dans le **DSL**, l'apprentissage ne se cantonne pas à une phase d'entraînement figée ; il se présente comme un **processus d'auto-organisation** continu, actif tant que le réseau demeure en service. Chaque entité peut ainsi **affiner** ou **redéfinir** ses connexions et son rôle lorsqu'affluent de nouvelles données, ou même de nouvelles entités.

Contrairement à un réseau neuronal classique, la **structure** du réseau n'est pas figée. Les pondérations synergiques $\omega_{i,j}(t)$ peuvent **croître** ou **décroître** jusqu'à atteindre un **seuil** de rupture ω_{\min} , entraînant ainsi la disparition effective de la liaison entre \mathcal{E}_i et \mathcal{E}_j . Ce mécanisme libère des ressources et prévient l'encombrement du réseau. À l'inverse, si une synergie autrefois négligeable devient significative, une **nouvelle liaison** peut se former pour consolider la coopération entre entités (voir la section 1.2.6 concernant les notations).

L'environnement ou la distribution des données peut également **fluctuer**. L'augmentation du bruit, l'apparition de nouvelles classes ou la transformation du contexte modifient les conditions d'apprentissage et d'adaptation du réseau. Dans un schéma statique, il faudrait réentraîner ou procéder à un "fine-tuning" coûteux. Dans le DSL, au contraire, l'**adaptation** est **naturellement** assurée par la mise à jour permanente des **liens synergiques** :

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta[S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)].$$

Si les informations nouvellement introduites modifient la synergie entre entités, la **topologie** du réseau se **recompose** spontanément, sans nécessité de réapprentissage global.

1.3.5.2. Formalisation mathématique de l'évolution dynamique

On peut représenter le réseau synergique à l'instant t par le graphe $G(t)$, constitué de nœuds $\{\mathcal{E}_1, \dots, \mathcal{E}_n\}$ et d'arêtes pondérées $\omega_{i,j}(t)$. L'évolution dans le temps se décrit par :

$$G(t+1) = \mathcal{F}(G(t), \mathcal{D}(t), \Theta),$$

où $\mathcal{D}(t)$ désigne les **données reçues** à l'instant t (ou sur la fenêtre $[t, t + \Delta]$), et Θ englobe les paramètres de mise à jour (taux η , régularisation τ , etc.).

Chaque entité \mathcal{E}_i peut maintenir un **état** $\mathbf{s}_i(t)$ permettant d'intégrer une forme de **mémoire** (cf. 1.2.1). Cet état influe sur la synergie :

$$S(\mathcal{E}_i, \mathcal{E}_j) = f(\mathbf{s}_i(t), \mathbf{s}_j(t), \mathcal{D}(t), \dots).$$

On a donc un **système dynamique couplé**, dans lequel la mise à jour des connexions $\omega_{i,j}(t)$ et celle des états $\mathbf{s}_i(t)$ interagissent :

$$\begin{cases} \omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)], \\ \mathbf{s}_i(t+1) = \mathbf{s}_i(t) + \gamma g(\mathbf{s}_i(t), \{\omega_{i,k}(t)\}, \dots). \end{cases}$$

On obtient alors un **couplage** entités-réseau qui peut donner naissance à des attracteurs, des cycles limites ou même des phénomènes de bifurcation.

Au-delà de la simple création/suppression de liens, il est possible que deux entités \mathcal{E}_a et \mathcal{E}_b **fusionnent** si leur collaboration est très constante, formant une **nouvelle entité** \mathcal{E}_{ab} . De même, une entité fortement hétérogène peut se **scinder** en plusieurs entités spécialisées si cela améliore la synergie globale. Ces mécanismes se décrivent par des règles de type :

$$\text{Fusion}(\mathcal{E}_a, \mathcal{E}_b) \Leftrightarrow \sum_{t'=t_0}^{t_1} \omega_{a,b}(t') > \delta_{\text{fusion}},$$

avec δ_{fusion} un seuil de stabilité sur une durée $[t_0, t_1]$. De tels phénomènes rappellent les “assemblées neuronales” durables ou la “spéciation” dans un contexte évolutionniste.

1.3.5.3. Avantages concrets d’une évolution en continu

De nombreux problèmes réels (prévision financière, capteurs industriels, suivi de l’activité cérébrale, etc.) impliquent des distributions de données qui **changent** dans le temps. Le DSL, grâce à son **évolution dynamique**, s’ajuste progressivement plutôt que de s’en tenir à un modèle figé.

Dans un réseau traditionnel, apprendre une nouvelle tâche (ou un nouveau domaine) peut faire oublier les acquis précédents, faute de mécanisme de préservation structurelle. Dans le **DSL**, il n’est pas nécessaire de redémarrer l’optimisation. Les **clusters utiles** peuvent **persister**, tandis que de nouvelles entités se connectent et ajustent la synergie pour assimiler un contexte différent.

Au fur et à mesure de l’évolution, si certaines liaisons n’apportent plus de gain, elles s’effacent ou diminuent. Ce phénomène **régule** la complexité du réseau en évitant une explosion exponentielle du nombre de liens. On peut donc voir le DSL comme un réseau à **topologie parcimonieuse** autorégulée.

1.3.5.4. Illustrations par l’exemple

Un **système DSL** qui reçoit en continu des mesures (température, pression, vibrations) peut servir à détecter des anomalies ou à anticiper des pannes. Lorsque les conditions d’usage évoluent progressivement, comme l’usure ou la modification de la chaîne de production, le réseau **réorganise** ses liens synergiques. Certains capteurs établissent de nouvelles complémentarités, tandis que des connexions historiques perdent de leur pertinence et s’affaiblissent. Cette **auto-organisation** permet de maintenir une alerte fiable, sans qu’il soit nécessaire de procéder à un nouvel entraînement global à intervalles réguliers.

Un **chatbot multimodal** traitant à la fois la voix, le texte et l’image du visage (via webcam) illustre également l’adaptabilité d’un réseau évolutif. À mesure que l’utilisateur se familiarise avec le

système (introduction de nouveaux termes, évolution de ses habitudes gestuelles), les entités “voix” et “vision” ajustent leurs connexions et peuvent, le cas échéant, **former un cluster** spécialisé pour reconnaître l'utilisateur concerné. Les anciennes tâches (reconnaître d'autres utilisateurs) ne sont pas oubliées, car elles persistent dans des clusters distincts ; la connaissance acquise est ainsi **préservée** sur le long terme.

1.3.5.5. Limites et défis de l'évolution continue

Si le réseau cherche à suivre de trop près les **fluctuations** de la synergie, il peut devenir **instable**, oscillant sans cesse ou détruisant prématurément des liens pertinents. Les paramètres η , τ , δ_{fusion} , etc., exigent donc un **calibrage** soigné pour maintenir un **équilibre** entre plasticité et stabilité (c'est le “dilemme stabilité–plasticité”).

Par ailleurs, dans un **modèle statique**, on fige les poids après l'apprentissage et on évalue la performance sur un ensemble de test. Dans le cadre d'un **réseau évolutif**, la performance est **susceptible de varier** au fil du temps, rendant nécessaire l'adoption de **métriques d'apprentissage continu** (p. ex. une mesure de l'erreur sur une fenêtre glissante).

Enfin, la reconfiguration permanente des liens implique un **coût de calcul**. Si le réseau compte n entités et que la synergie est évaluée entre toutes les paires, on fait face à $O(n^2)$ opérations. Des stratégies de **parsimonie** (seuil de rupture, mise à jour partielle, échantillonnage) sont souvent indispensables pour demeurer **scalable** dans des systèmes de grande ampleur.

1.3.5.6. Conclusion : une architecture “vivante” pour l'IA

L'**évolution dynamique** et l'**adaptation en continu** font du **Deep Synergy Learning** une forme de réseau “vivante”, capable de remodeler son organisation interne au fur et à mesure que l'environnement (ou les données) se transforment. Ce **changement de paradigme** – de l'entraînement ponctuel et figé à une auto-organisation permanente – ouvre la voie à des **applications** de l'IA plus flexibles et plus robustes, réduisant la dépendance à la supervision humaine et accroissant la **longévité** des modèles dans des scénarios réels et évolutifs.

Les sections suivantes, **1.3.6** et **1.3.7**, approfondiront deux aspects essentiels de cette approche. La section **1.3.6** explorera la **robustesse** et la **résilience** apportées par l'auto-organisation, tandis que la section **1.3.7** abordera les **pistes** vers une approche plus générale de l'apprentissage, potentiellement orientée vers une IA plus proche de la **cognition** ou de l'**intelligence générale**.

1.3.6. Impacts sur la Robustesse et la Résilience des Modèles

L'une des conséquences les plus marquantes de l'**auto-organisation** et de la **synergie informationnelle** dans le **Deep Synergy Learning (DSL)** réside dans l'**amélioration** de la **robustesse** et de la **résilience** des systèmes d'apprentissage. Contrairement aux approches classiques, souvent vulnérables aux perturbations, aux changements de distribution ou aux défaillances partielles, le DSL tire parti de sa structure **évolutive** et **coopérative** pour mieux

absorber ces aléas. Dans la présente section, nous examinons les notions de robustesse et de résilience, puis montrons comment elles s’expriment dans un réseau synergique.

1.3.6.1. Définitions et enjeux

La **robustesse** d’un modèle désigne sa capacité à **maintenir** un niveau de performance élevé malgré la présence de **perturbations** ou d’**incertitudes**. Dans le contexte de l’IA, ces perturbations peuvent inclure :

- Du **bruit** dans les données (capteurs défectueux, images floues, signaux audio corrompus, etc.).
- Des **attaques adversariales** (petites perturbations “adversarial noise” rendant le modèle confus).
- Des **incohérences** ou **lacunes** dans les échantillons (valeurs manquantes, distribution très variée).

Un système robuste parvient à faire face à ces dégradations sans “collapsus” brutal de la performance.

La **résilience** va plus loin que la robustesse en impliquant non seulement la résistance aux perturbations, mais aussi la **capacité à récupérer** ou à **se réorganiser** après un choc. En IA, cela signifie que si la distribution des données change, ou si le système subit une panne partielle (certains capteurs tombent en panne, certains flux de données disparaissent), la résilience se manifeste par la **reconfiguration** interne permettant de continuer la tâche ou de s’en rapprocher au mieux.

1.3.6.2. Mécanismes de robustesse dans un réseau synergique

Dans le DSL, plusieurs entités d’information peuvent recouvrir partiellement la même “zone de compétence” ou la même modalité, tout en se distinguant suffisamment pour apporter une **valeur ajoutée**. Du point de vue de la robustesse, cela signifie qu’un flux corrompu (ou une entité dysfonctionnelle) n’est pas catastrophique, car d’autres entités peuvent **prendre le relais**.

$$\omega_{i,j}(t) \rightarrow \omega_{k,j}(t+1), \quad \text{si } \mathcal{E}_i \text{ défaille, alors } \mathcal{E}_k \text{ peut compenser.}$$

Lorsque la synergie entre une entité défectueuse \mathcal{E}_{def} et les autres entités chute (bruit, erreur répétée, etc.), les pondérations $\omega_{\text{def},j}(t)$ s’affaiblissent progressivement. Le réseau “apprend” ainsi à **alléger** la connexion avec la source fautive et à **rediriger** l’information vers des entités plus fiables. Ce mécanisme évite un fort impact d’une entité isolée dysfonctionnelle.

Les entités réagissent localement au **bruit** ou aux **incohérences** en resserrant leurs liens synergiques avec celles qui restent cohérentes. On observe alors la création de **clusters** (voir 1.2.6) qui s’auto-assemblent autour de flux fiables ou complémentaires, renforçant la robustesse globale.

$$\mathcal{C}^* = \arg \max_{\mathcal{C}} \sum_{(i,j) \in \mathcal{C} \times \mathcal{C}} \omega_{i,j}, \quad \text{en excluant } \mathcal{E}_{\text{def}}.$$

1.3.6.3. Résilience via l'adaptation en continu

Le DSL met en œuvre un apprentissage **continu** (section 1.3.5), ce qui signifie que le réseau s'ajuste aux nouvelles données ou aux nouveaux contextes sans nécessiter une refonte globale. Ainsi, si une modalité devient soudainement imprécise (caméra saturée de lumière, micro exposé à un fort bruit ambiant), le réseau réorganise ses **pondérations synergiques** pour ne plus dépendre de cette source.

$$\omega_{\text{camera}, \mathcal{E}_j}(t+1) = \omega_{\text{camera}, \mathcal{E}_j}(t) - \eta \left[\tau \omega_{\text{camera}, \mathcal{E}_j}(t) \right] \quad \text{si } S(\mathcal{E}_{\text{camera}}, \mathcal{E}_j) \approx 0 \text{ (bruit élevé)}.$$

La résilience suppose la **recupération** ou le maintien d'une performance acceptable même après un choc. Le fait de **supprimer** ou d'**affaiblir** les liens inutiles (ou trompeurs) et de **renforcer** les liens pertinents génère une reconfiguration topologique :

$$G(t+1) = \mathcal{U}[G(t), \{\omega_{i,j}\}, \text{nouveaux contextes}].$$

Ainsi, le réseau peut se redessiner sans qu'on doive recourir à une procédure hors-ligne complexe.

Les entités (et clusters) qui étaient **efficaces** pour d'anciennes tâches peuvent persister en parallèle, assurant une forme de **mémoire** ou de **transfert** vers des scénarios futurs. Cette coexistence entre anciens et nouveaux liens renforce la **résilience** du réseau. Plutôt que d'éliminer systématiquement ses anciennes connaissances, il les conserve tant qu'elles restent **synergiques**, évitant ainsi le phénomène de **catastrophic forgetting**.

1.3.6.4. Exemples d'Applications Robustes et Résilientes

Le **Deep Synergy Learning (DSL)** se révèle particulièrement adapté aux situations où la robustesse et la résilience constituent des critères essentiels. Deux scénarios illustrent la manière dont la **synergie informationnelle** et la **dynamique auto-organisée** du DSL permettent d'absorber des pannes, de réagir à des anomalies et de préserver la continuité opérationnelle. Dans chaque cas, la structuration coopérative du réseau offre une forme de **tolérance aux défaillances**, car l'affaiblissement d'un module ou d'un capteur ne remet pas en cause l'ensemble de la configuration. Les **exemples** ci-dessous mettent en évidence ces mécanismes de résilience et soulignent les **bienfaits** d'un réseau qui se **reconfigure** en temps réel.

A. Robotique et Systèmes Autonomes

Considérons un **robot** doté de multiples capteurs, tels que des caméras, un LIDAR, des gyroscopes, ainsi que d'autres senseurs spécialisés. Dans ce contexte, un modèle DSL peut absorber la défaillance partielle ou totale d'un capteur, sans qu'il soit nécessaire de procéder à un réapprentissage global. Cette propriété découle du caractère **auto-adaptatif** du Synergistic Connection Network (SCN) sous-jacent, où chaque capteur \mathcal{C}_i est traité comme une **entité** de

l'ensemble, tandis que les liens $\omega_{i,j}$ mesurent la **coopération** ou la **complémentarité** de deux capteurs i et j .

Dans le cas d'une **défaillance**, supposons que le capteur \mathcal{C}_k soit victime d'un bruit extrême ou d'une panne soudaine. Le **score de synergie** $S(\mathcal{C}_k, \mathcal{C}_m)$ avec les autres capteurs \mathcal{C}_m diminuerait alors, ce qui provoquerait une mise à jour négative des liens $\omega_{k,m}$ par l'équation

$$\omega_{k,m}(t+1) = \omega_{k,m}(t) + \eta[S(\mathcal{C}_k, \mathcal{C}_m) - \tau \omega_{k,m}(t)],$$

où η désigne le taux d'apprentissage local et τ le coefficient de régulation. À mesure que la **synergie** s'affaiblit, la pondération $\omega_{k,m}$ décroît, de sorte que ce capteur devient quasi inactif au sein du SCN. Les autres capteurs, préservant des liens ω élevés entre eux s'ils demeurent fiables et cohérents, continuent de coopérer pour fournir les informations nécessaires à la navigation ou à la réalisation des tâches robotiques. L'ajustement s'effectue en continu, sans qu'un entraînement complet soit relancé, montrant ainsi la **résilience** inhérente au DSL.

B. Surveillance Critique

Dans les domaines de la **surveillance** ou du **monitoring** (installations sensibles, applications médicales, détection d'intrusions), il est fréquent de recourir à un grand nombre de capteurs hétérogènes. Un système basé sur le DSL peut y déployer ses mécanismes d'**auto-organisation** de manière à détecter plus vite d'éventuelles anomalies et à tolérer les dysfonctionnements partiels d'un ou plusieurs capteurs. Le déploiement s'opère suivant la logique où chaque capteur \mathcal{D}_i constitue un **nœud** du SCN, et où le **score de synergie** $S(\mathcal{D}_i, \mathcal{D}_j)$ quantifie la correspondance ou la complémentarité entre deux capteurs.

Lorsque l'un de ces nœuds présente une anomalie ou se déconnecte, la **pondération** $\omega_{i,j}$ chute au fil des itérations, car la synergie mesurée devient faible. Mathématiquement, le processus de mise à jour

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta[S(\mathcal{D}_i, \mathcal{D}_j) - \tau \omega_{i,j}(t)]$$

réduit progressivement les liens inutilisables ou compromis. La structure globale demeure néanmoins cohérente, car les capteurs restant fiables maintiennent des valeurs de ω importantes, ce qui préserve la capacité du réseau à **fonctionner**.

Le phénomène inverse se produit lorsqu'une alarme est déclenchée de façon synchronisée entre plusieurs capteurs. Dès lors que ceux-ci fournissent un signal convergent sur la présence d'un événement inhabituel, leur **synergie** augmente, et ils forment un **cluster** spécialisé dans l'analyse d'un risque potentiel. Cette coopérative rapide et flexible permet au SCN de hiérarchiser l'information critique. L'avantage repose sur l'**absence de réapprentissage intégral** et sur la capacité du système à s'**auto-réorganiser** pour mettre en avant les flux les plus pertinents.

Conclusion

Que ce soit dans un scénario de **robotique** multi-capteurs ou dans un dispositif de **surveillance** critique, le **Deep Synergy Learning** démontre ainsi son **aptitude** à gérer la panne ou la perte de fiabilité d'un sous-module sans perturber l'ensemble de l'architecture. La dynamique de **mise à jour** des synergies et des liens ω assure une continuité de service et un renforcement rapide des

signaux valides, ce qui illustre la solidité opérationnelle et l'efficacité de cette **approche auto-organisée** pour la gestion des situations critiques ou complexes.

1.3.7. Perspectives pour une Approche plus Globale de l'Apprentissage

Les sections précédentes (1.3.1 à 1.3.6) ont mis en évidence l'importance de l'**auto-organisation** dans le Deep Synergy Learning (DSL), ses fondements biologiques et cognitifs (1.3.1), les concepts d'émergence et de feedback (1.3.2), la comparaison avec les méthodes classiques (1.3.3), le rôle crucial de la multi-modalité (1.3.4), la dynamique adaptative (1.3.5) et l'impact sur la robustesse/résilience (1.3.6). Nous abordons maintenant une question centrale. **Dans quelle mesure** cette architecture auto-organisée peut-elle contribuer à une **approche plus globale** de l'apprentissage, ouvrant la voie à une IA plus générale, moins cloisonnée et plus proche de la **cognition humaine** ?

1.3.7.1. Vers une intégration de multiples paradigmes

L'un des attraits majeurs du DSL réside dans sa **flexibilité** structurelle, permettant l'intégration simultanée de tâches variées (reconnaissance d'images, traitement du langage, analyse de signaux). Plutôt que de multiplier les sous-modules indépendants, le DSL **mutualise** la capacité d'adaptation et de co-évolution. Cela ouvre la voie à des systèmes capables d'aborder des **problèmes multi-domaines** de façon unifiée.

Beaucoup de chercheurs s'interrogent sur la possibilité de combiner une **représentation symbolique** (logique, règles, ontologies) et des **réseaux neuronaux** (approche sub-symbolique). Le DSL offre un espace où des **entités symboliques** (représentant des concepts, des règles) pourraient coexister avec des **entités sub-symboliques** (features non supervisées, clusters contextuels), et forger des **liens synergiques** si ces représentations s'avèrent mutuellement bénéfiques.

$$\omega_{\text{symbolique,sub-symbolique}}(t+1) = \omega_{\text{symbolique,sub-symbolique}}(t)\eta[S(\mathcal{E}_{\text{symbolique}}, \mathcal{E}_{\text{sub-symbolique}})].$$

Une autre piste consiste à inclure des **boucles de récompense** (façon apprentissage par renforcement) à l'intérieur du DSL, de manière que la **synergie** prenne en compte non seulement l'interaction entre entités, mais aussi un **signal de performance** plus global. Cela pourrait mener à des architectures où l'**exploration** et la **sélection** des connexions synergiques s'effectuent au service d'une stratégie d'agent, par exemple dans un environnement en évolution (robotique, jeux, etc.).

1.3.7.2. Apprentissage contextuel et raisonnement adaptatif

Dans de nombreux problèmes (analyse de scène, dialogue, diagnostic médical), l'**environnement** ou la **situation** évolue. Le DSL permet de faire émerger des **clusters contextuels**, regroupant les entités les plus pertinentes pour un contexte donné. Ces clusters peuvent ensuite se **désagréger** ou

se **recombina** quand le contexte change, offrant un niveau de **contextualisation dynamique** qu'on retrouve rarement dans les modèles hiérarchiques figés.

Plutôt que de raisonner par règle (“si A alors B”), le réseau peut **découvrir** que certaines combinaisons d’entités $\{\mathcal{E}_i, \mathcal{E}_j, \dots\}$ produisent un **effet global** pertinent. On pourrait formaliser cette découverte comme un **raisonnement émergent**, où l’apparition d’un cluster synergique équivaut à la création d’un “concept” ou d’une “hypothèse” confirmée par l’amélioration de la performance ou de la cohérence.

$$\text{Concept}_\alpha \leftrightarrow \{\mathcal{E}_i, \mathcal{E}_j, \dots\} \quad \text{avec} \quad \sum_{(i,j) \in \alpha} \omega_{i,j} > \theta.$$

Les entités dans le DSL peuvent porter un **état interne** (1.2.1), ce qui permet de **mémoriser** certains événements, associations ou transitions. Couplé à l’idée de clusters, on obtient une forme de mémoire **distribuée** et **adaptative**, capable de se ré-agencer au fil du temps, au lieu d’être confinée dans des architectures rigides (ex. LSTM).

1.3.7.3. Ouverture vers une IA plus proche de la cognition biologique

Nous avons déjà souligné les parallèles entre la synapse biologique et les connexions synergiques (sections 1.3.1, 1.3.2). Un système auto-organisé qui évolue au fil du temps, fusionne ou dissocie des entités, renforce ou supprime des liens, **s’approche** d’une **dynamique neuronale** élémentaire. Il est donc envisageable que des **architectures DSL** se montrent plus aptes à **simuler** ou **comprendre** certains phénomènes cognitifs.

Bien que spéculative, la question d’une **conscience artificielle** ou d’une **auto-consistance** cognitive pourrait trouver dans le DSL un champ d’exploration. Certains travaux suggèrent qu’un degré d’**intégration d’information** (comme l’approche de la “Phi measure” de Giulio Tononi) est requis pour émerger une forme de conscience. Si le DSL parvient à développer des **clusters** hautement connectés et persistants (au sens “informationnellement intégrés”), il pourrait servir de terrain expérimental pour avancer sur ces hypothèses.

L’IA forte (IAG) ou Intelligence Artificielle Générale se définit par la capacité d’un système à maîtriser un large éventail de tâches, à apprendre de manière autonome et à faire face à des environnements variés. La **plasticité** et la **co-évolution** du DSL, sa manière de laisser chaque entité s’adapter, fusionner, se spécialiser, suggèrent un potentiel pour **surmonter** les obstacles auxquels se heurtent les méthodes rigides (catastrophic forgetting, absence de transfert, etc.). Cela ne garantit pas l’émergence d’une IAG, mais en offre une piste conceptuelle plus proche du fonctionnement adaptatif du vivant.

1.3.7.4. Principaux défis et voies de recherche

Le **Deep Synergy Learning (DSL)** introduit un cadre inédit, reposant sur une **interaction dynamique** et une **auto-organisation** continue des entités d’information. Malgré l’attrait de cette proposition, il subsiste plusieurs défis mathématiques, algorithmiques et pratiques, dont la résolution s’avère essentielle pour la généralisation à grande échelle. Sur le plan **computationnel**,

l'évaluation des synergies entre un nombre important d'entités peut impliquer une complexité en $O(n^2)$, où n désigne le nombre d'entités. Cette explosion combinatoire requiert des stratégies de **parsimonie** et de **sparsification** afin de restreindre la croissance du nombre de connexions à évaluer. Des techniques d'**échantillonnage** adaptatif peuvent aussi être envisagées, de manière à ne considérer que les paires d'entités présentant une synergie potentiellement élevée. D'un point de vue plus formel, on peut réduire la densité des connexions actives en imposant un mécanisme de seuil sur les poids ω_{ij} , de sorte qu'une liaison (i, j) ne soit conservée que si

$$\omega_{ij} > \omega_{\min},$$

avec ω_{\min} un paramètre choisi pour garantir la structure globale du réseau.

S'agissant de l'**interprétabilité** et de l'**explicabilité**, la capacité du DSL à faire émerger des **clusters** et à redessiner en continu ses connexions favorise une organisation plus fluide, mais rend ardue l'analyse a posteriori des décisions. Pour éclairer la nature d'une prédiction, il devient nécessaire de retracer l'historique des synergies et d'examiner l'état interne des entités concernées. La mise à jour de chaque poids ω_{ij} peut être décrite par une équation de type

$$\omega_{ij}(t+1) = \omega_{ij}(t) + \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{ij}(t)],$$

où η désigne le taux d'apprentissage et τ le coefficient de décroissance. Cette dynamique rend possible l'apparition ou la suppression de liens dont la contribution n'est plus jugée pertinente, tout en renforçant ceux qui semblent les plus profitables au regard de la synergie mesurée. Cependant, un tel fonctionnement complique la **traçabilité**. Pour extraire des règles ou des justifications, il serait nécessaire de stocker ou de synthétiser l'ensemble des états successifs ainsi que les évaluations de $S(\mathcal{E}_i, \mathcal{E}_j)$.

La question de la **convergence** suscite également un vif intérêt. L'étude de la stabilité et des attracteurs potentiels liés aux mises à jour des poids dans le cadre d'un **système dynamique** exige d'analyser l'évolution de l'**énergie** du réseau ou d'une **fonction de Lyapunov** associée. Si l'on considère une énergie

$$E(t) = - \sum_{i \neq j} \omega_{ij}(t) S(\mathcal{E}_i, \mathcal{E}_j),$$

la décroissance de $E(t)$ au fil des itérations peut servir d'indicateur de convergence vers un état stable où les pondérations cessent d'évoluer de manière significative. L'identification de conditions garantissant l'existence d'un minimum global ou local, ainsi que la caractérisation de la vitesse de convergence, demeurent des questions ouvertes pour la communauté scientifique. La présence de boucles de rétroaction complexes, de reconfigurations topologiques et d'interactions non linéaires peut engendrer des comportements oscillatoires, voire chaotiques, si aucune **régularisation** adéquate n'est mise en place.

Dans les **domaines critiques** (santé, finance, transport, sécurité), le DSL pourrait apporter une meilleure **robustesse** grâce à la **fusion multimodale** et à la **répartition adaptative** des informations, permettant de gérer des données hétérogènes ou partiellement manquantes. Toutefois, de telles applications exigent des garanties fortes sur la fiabilité et la sûreté de fonctionnement, impliquant des **processus de certification** exigeants. Il devient essentiel d'y

intégrer des mécanismes de **sécurité** contre d'éventuelles attaques adversariales, où de subtiles perturbations de l'entrée peuvent tromper même des systèmes neuronaux performants. La dynamique adaptative et coopérative du DSL pourrait, d'un côté, renforcer la **résilience** en détectant des incohérences via les synergies internes, mais de l'autre côté, elle complique l'établissement de **bornes** de sûreté ou de **preuves** formelles de robustesse.

Ainsi, malgré un potentiel considérable, le DSL fait face à des enjeux théoriques et pratiques déterminants pour son adoption à grande échelle. La **réduction de la complexité algorithmique**, l'**amélioration de l'interprétabilité**, l'**étude minutieuse de la convergence** et l'**adaptation aux environnements critiques** constituent autant de pistes de recherche centrales pour consolider ce paradigme. Les avancées futures consisteront notamment à concevoir des **stratégies d'approximation** limitées en ressources de calcul, à développer des **méthodes d'exploration** et de **visualisation** capables d'exposer la structure émergente du réseau, et à établir des **cadres mathématiques** plus complets validant la stabilité et la fiabilité du processus d'**auto-organisation**. L'ensemble de ces travaux pourra contribuer à faire du Deep Synergy Learning un jalon clé dans l'évolution de l'IA vers une plus grande **plasticité**, **résilience** et **intelligence**.

1.3.7.5. Conclusion : vers un nouveau paradigme d'apprentissage

En somme, l'auto-organisation et la synergie, éléments centraux du **Deep Synergy Learning**, ouvrent la voie à une **approche plus globale** de l'apprentissage, susceptible d'englober plusieurs paradigmes existants (supervisé, non supervisé, renforcement, symbolique, etc.) tout en proposant des mécanismes d'évolution continue, de raisonnement contextuel et de mémorisation distribuée.

Cette vision ne constitue pas encore un aboutissement. De nombreux **défis** scientifiques, méthodologiques et éthiques, tels que la **scalabilité**, l'**interprétabilité** et la **robustesse**, restent à relever. Néanmoins, le DSL fournit d'ores et déjà un **cadre unificateur** pour penser l'**IA adaptative**, apte à accueillir de multiples flux d'information et à co-évoluer de façon organique, en s'inspirant des systèmes vivants.

Les chapitres suivants approfondiront encore la **formulation mathématique** du DSL, ses **algorithmes** et ses **applications pratiques**, en vue de concrétiser ce nouveau paradigme et d'explorer son potentiel pour des systèmes intelligents, au-delà des limites imposées par les architectures hiérarchiques classiques.

1.4. Architecture Générale du DSL

Dans les chapitres précédents, nous avons défini les **fondements conceptuels** du Deep Synergy Learning (DSL), en soulignant notamment l'importance de l'**auto-organisation** et des **synergies** entre entités d'information (sections 1.2 et 1.3). Nous savons désormais que chaque entité ne se limite pas à un simple vecteur de données. Elle peut **évoluer**, **interagir** et **générer de nouvelles représentations** lorsqu'elle coopère avec d'autres entités. Ces processus de coopération ne sont pas imposés par un schéma hiérarchique rigide, mais émergent librement à mesure que les entités détectent un **gain** dans leur association.

Afin de concrétiser ces principes dans un **cadre unifié**, il est nécessaire de décrire l'**architecture générale** du DSL. Autrement dit, comment ces entités s'organisent-elles ? Quelles structures se dégagent quand la synergie est élevée entre certaines paires (ou ensembles) d'entités ? Et comment modéliser leur évolution au fil du temps ? La section 1.4 répond à ces questions en introduisant :

- Les **principes de base** des entités et des liens (1.4.1)
- La notion de **Synergistic Connection Network (SCN)** (1.4.2)
- La formation de **clusters** et de **macro-clusters** (1.4.3)
- Les différentes **fonctions de synergie** (1.4.4)
- Le fonctionnement **adaptatif** des pondérations (1.4.5)
- Les interactions **directes** et **indirectes** (1.4.6)
- Les cas particuliers de **synergie binaire** et **n-aire** (1.4.7)

C'est cette **architecture globale** du **DSL** qui confère au système sa **plasticité** et sa **capacité d'adaptation**. Plutôt que de transmettre passivement les données à travers des couches fixes, chaque entité peut, en temps réel, **renforcer ou affaiblir** ses liens avec d'autres, **former des clusters spontanés**, et **faire émerger** des représentations plus riches que la simple somme de leurs composantes individuelles.

1.4.1. Principe de Base : Entités et Liens Synergiques

Le **Deep Synergy Learning** envisage chaque **entité d'information** comme un **nœud actif** de l'architecture globale. Plutôt que de propager passivement un signal, comme le ferait un neurone dans un réseau hiérarchique, cette entité adopte un rôle décisif en disposant de **paramètres internes** θ_k , souvent constitués d'une **représentation** \mathbf{x}_k (vectorielle ou tensorielle), d'un **état interne** \mathbf{s}_k qui se modifie au gré des interactions, et de tout **hyperparamètre** requis pour la modalité concernée (par exemple un encodeur audio ou un embedding textuel). Chaque entité s'accompagne également de **mécanismes de décision** qui lui permettent d'**observer** la synergie obtenue lorsqu'elle se connecte à d'autres entités et, selon cette observation, de renforcer, d'affaiblir ou de rompre certains liens. Elle peut même, dans les cas où la coopération se révèle durablement élevée, envisager une **fusion** avec une autre entité ou, au contraire, une **spécialisation** en segmentant ses propres paramètres pour répondre à des tâches spécifiques. Cette description,

encore préliminaire, fait déjà ressortir l'idée que les **nœuds** du DSL sont bien plus dynamiques que dans un réseau de neurones classique, puisque chacun "recherche" les coopérations les plus fructueuses pour améliorer sa représentation tout en contribuant à la robustesse du réseau dans son ensemble.

Chaque entité \mathcal{E}_i peut se connecter à une autre entité \mathcal{E}_j par le biais d'une **pondération synergique** $\omega_{i,j}(t)$, laquelle traduit la **valeur ajoutée** perçue dans cette association. Quand la synergie $S(\mathcal{E}_i, \mathcal{E}_j)$ est forte, on observe un accroissement de $\omega_{i,j}(t)$, tandis que dans le cas contraire, ce lien tend à se résorber puis à disparaître. L'évolution de $\omega_{i,j}$ se modélise alors par une équation simple, par exemple

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)],$$

où η désigne le taux d'apprentissage et τ un paramètre de régularisation. Cette formulation, exposée plus en profondeur dans les sections ultérieures, peut aussi s'enrichir de seuils pour limiter la croissance des poids ou supprimer les connexions trop faibles, assurant de la sorte une **régulation** de la densité du graphe. Une fois que l'on modélise ainsi les **liens synergiques**, on se dote d'une **matrice d'adjacence** évolutive qui dicte quels nœuds interagissent, quand ils le font et avec quelle intensité.

La grande différence avec une organisation hiérarchique est que l'information n'est plus obligée de transiter selon des voies prédéfinies. Toute entité peut potentiellement échanger avec n'importe quelle autre, pour peu que la **pondération synergique** atteigne un niveau jugé satisfaisant. Dans le cadre d'une tâche multimodale, il est donc envisageable que des entités "visuelles" se lient directement à des entités "textuelles" si elles y trouvent un bénéfice mutuel. Il se forme alors des **clusters** éphémères ou plus durables, réunissant deux, trois ou davantage d'entités, et la persistance de ces regroupements dépend de la permanence de leur synergie. De tels micro-réseaux se constituent et se défont librement, donnant lieu à une **auto-organisation** dont l'architecture n'a pas été spécifiée a priori mais découle de la dynamique interne du système.

Cette perspective se décrit volontiers comme un **graphe évolutif** $G(t)$. Les **nœuds** correspondent aux entités $\{\mathcal{E}_1, \dots, \mathcal{E}_n\}$, et les **arêtes** représentent les liaisons $\omega_{i,j}(t)$. À chaque itération, une règle d'actualisation \mathcal{U} recalcule $\omega_{i,j}(t+1)$ en tenant compte de la **synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$ et de la **régularisation** (qui inclut, au besoin, la suppression des liens faibles). Cela rend possibles des analyses croisées avec la **théorie des graphes** ou la **dynamique non linéaire**, et l'on peut même faire appel à des méthodologies d'**optimisation combinatoire** lorsqu'on cherche, par exemple, à minimiser un coût global ou à identifier une configuration de connexions correspondant à un optimum local.

Cette approche basée sur des **nœuds actifs** et des **arêtes dynamiques** confère au **DSL** une **flexibilité** qui se révèle supérieure à celle de nombreuses architectures traditionnelles. La topologie n'est plus figée dans des couches de traitement ; elle se réorganise au gré des opportunités de synergie détectées, ce qui est d'un intérêt crucial pour les situations multimodales ou pour l'adaptation à des données en évolution. Les entités elles-mêmes peuvent fusionner, se subdiviser ou changer de paramétrage, et toute la structure s'oriente vers une organisation où la synergie la plus forte est mise en avant. Cette logique favorise l'apparition de **propriétés émergentes** comme la consolidation de **macro-clusters** ou l'élaboration de schémas coopératifs inédits, sans qu'aucune hiérarchie ne soit imposée dès le départ.

En somme, la clef de voûte du **DSL** réside dans l'idée que chaque entité constitue un **nœud autonome**, et chaque connexion, un **lien synergique** dont l'évolution est régie par le degré d'enrichissement mutuel. Les observations de clusters, de reconnections ou de regroupements imprévus y sont donc la norme, et ce caractère **distribué** et **adaptatif** positionne le DSL comme un prolongement original des paradigmes existants, taillé pour gérer la variété et la complexité croissantes des données actuelles.

1.4.2. Présentation du Synergistic Connection Network (SCN)

Le **Synergistic Connection Network (SCN)** constitue la **pièce maîtresse** du Deep Synergy Learning (DSL) en matérialisant à la fois l'**espace** dans lequel les entités évoluent et le **mécanisme** même qui autorise la naissance, la transformation ou la disparition de leurs liens. Alors que les réseaux de neurones classiques reposent sur une **topologie fixée** avant l'entraînement, le SCN propose, au contraire, une **structure adaptive** dont les entités et leurs connexions se reconfigurent de manière autonome au fil du temps.

La représentation du **SCN** sous forme de **graphe** positionne les **entités** $\{\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_n\}$ en tant que **nœuds**, reliés par des **pondérations synergiques** $\omega_{i,j}$ évolutives. Chaque entité correspond à un bloc fonctionnel, qu'il soit **visuel**, **textuel** ou **auditif**, et ajuste dynamiquement ses connexions en fonction de la **synergie** détectée. Elle peut ainsi **renforcer** ou, au contraire, **réduire** sa liaison avec une autre entité selon la pertinence de leur interaction. Cette **plasticité** donne au **SCN** un caractère **vivant**. Sa structure n'est jamais figée mais se **recompose** en permanence, en conservant les connexions jugées utiles et en éliminant celles qui ne le sont plus.

A. Les principes fondateurs du SCN

Coopération locale, cohérence globale

Au cœur du **SCN**, chaque liaison $\omega_{i,j}$ évolue selon une **décision locale**. Deux entités évaluent leur **synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$, qu'elle repose sur une **similarité**, une **co-information** ou un **gain de performance**, puis ajustent la pondération $\omega_{i,j}$. Pourtant, cette dynamique locale influence la **structure globale**. En synchronisant leurs mises à jour, des milliers ou des millions de connexions peuvent, en quelques itérations, donner naissance à des **macro-structures** ou des **clusters** (voir section 1.4.3).

Adaptation continue

Plutôt qu'un apprentissage ponctuel, le SCN procède par **itérations** successives. Les pondérations $\omega_{i,j}$ se recalculent via une équation du type

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)],$$

où η indique le rythme de l'apprentissage et τ limite la persistance d'un lien. Cette **dynamique** empêche toute stagnation et confronte en permanence l'évolution des connexions aux signaux de synergie qui apparaissent ou disparaissent.

Mécanismes de régulation

Pour qu'un trop grand nombre de liaisons ne se maintienne artificiellement, on peut **couper** celles dont la pondération demeure en deçà d'un certain **seuil** ω_{\min} . À l'inverse, fixer un **plafond** ω_{\max} empêche l'explosion de liens exagérément forts. Ces stratégies renforcent la **parsimonie** du réseau. En ne conservant que les connexions bénéfiques, elles favorisent la formation de sous-ensembles stables et révèlent des motifs structurels plus riches.

B. Du réseau neuronal traditionnel au SCN

Un **réseau de neurones classique** repose sur une **architecture prédéfinie**. Chaque couche contient un nombre fixe de neurones, et les connexions entre elles restent invariables. Bien que les poids internes soient ajustés par la rétropropagation du gradient, la **topologie** du réseau, comprenant le nombre de couches, la disposition des neurones et le schéma des connexions, demeure **immuable**.

À l'inverse, le **SCN** adopte un paradigme fondamentalement différent. La **topologie évolue** spontanément, permettant aux connexions d'apparaître ou de disparaître en fonction de la **synergie** détectée. Les entités sont **actives** et disposent d'un état interne ou de mécanismes de représentation qui peuvent s'adapter pour optimiser leur **coopération**. Le réseau peut ainsi **se réinventer** en fonction de l'arrivée de nouvelles données ou de changements contextuels. Des **clusters spécialisés** émergent et disparaissent naturellement, sans nécessiter d'intervention extérieure pour redéfinir la structure.

Avec cette **plasticité**, le SCN parvient à incorporer et à privilégier les synergies révélées par les données, adaptant la répartition des connexions bien plus librement qu'un réseau hiérarchique prédéterminé.

C. Domaines d'application du SCN

Les **domaines d'application du SCN** exploitent sa capacité à s'adapter dynamiquement aux interactions entre entités et à optimiser l'apprentissage de manière continue.

La **multimodalité fluide** repose sur l'association flexible des différentes modalités, qu'il s'agisse de la vision, du texte, de l'audio ou de capteurs spécialisés. Le **SCN** établit et supprime des connexions en fonction de la **valeur ajoutée** qu'elles apportent. Les entités **visuelles** peuvent ainsi se connecter directement à des entités **textuelles** lorsqu'un **gain mutuel** est identifié, favorisant une intégration optimale des données.

L'**apprentissage continu et évolutif** permet au **SCN** de **réajuster** sa configuration en permanence dans un environnement changeant. Les liens utiles sont renforcés tandis que les connexions moins pertinentes sont progressivement délaissées. Cette plasticité itérative confère au réseau une **résilience** accrue, évitant ainsi un réapprentissage intégral à chaque modification de distribution des données.

La **découverte de patrons "n-aires"** va au-delà des interactions par paires. Le **SCN** favorise l'émergence de **macro-clusters**, où plusieurs entités $\{\mathcal{E}_1, \dots, \mathcal{E}_m\}$ coopèrent pour générer une **plus-value collective** qui dépasse la simple somme de leurs synergies binaires.

Dans la section suivante (1.4.3), nous verrons comment ces liens synergiques favorisent la naissance de **clusters auto-organisés**, et en quoi ceux-ci peuvent se regrouper en **macro-clusters**

pour donner naissance à des entités d'information plus puissantes, ou plus abstraites, dans la démarche d'apprentissage distribué du Deep Synergy Learning.

1.4.3. Notion de Cluster et de Macro-Cluster

Dans le **Synergistic Connection Network (SCN)**, décrit à la sous-section précédente (1.4.2), chaque entité d'information \mathcal{E}_i est reliée aux autres par des **pondérations synergiques** $\omega_{i,j}(t)$ qui évoluent dans le temps. Ce mécanisme permet l'**émergence spontanée** de sous-structures au sein du réseau. Des groupes d'entités présentant une **synergie élevée** ont tendance à **collaborer** plus fréquemment, formant ainsi des **clusters**.

Au-delà de ces regroupements de base, il est possible que **plusieurs** clusters se regroupent encore à un niveau supérieur pour former des **macro-clusters**. Cette section (1.4.3) approfondit la *définition mathématique* de ces concepts, leurs *règles d'apparition* et leurs *conséquences* sur la dynamique d'apprentissage dans le **Deep Synergy Learning (DSL)**.

Dans la perspective du **Synergistic Connection Network (SCN)**, il est souvent utile de modéliser la structure courante du réseau par un **graphe** $G(t)$ dont les **sommets** $\{\mathcal{E}_1, \dots, \mathcal{E}_n\}$ correspondent aux entités d'information, et dont chaque **arête** est dotée d'une **pondération synergique** $\omega_{i,j}(t)$. À l'instant t , ces pondérations se réunissent dans une **matrice** :

$$W(t) = [\omega_{i,j}(t)]_{1 \leq i,j \leq n}.$$

Cette représentation matricielle capture simultanément l'ensemble des liens du réseau, chaque $\omega_{i,j}(t)$ traduisant l'intensité de la coopération entre l'entité \mathcal{E}_i et l'entité \mathcal{E}_j . Dans un tel graphe, un **cluster** $\mathcal{C} \subset \{1, \dots, n\}$ (c'est-à-dire un sous-ensemble des indices d'entités) se caractérise par des **connexions internes** plus **denses** ou plus **fortes** que ses connexions avec l'extérieur. Plusieurs manières existent pour formaliser cette notion de cluster :

Un **premier** point de vue consiste à **maximiser la somme** des pondérations internes. Plus précisément, on cherche un sous-ensemble \mathcal{C} maximisant

$$\sum_{i \in \mathcal{C}, j \in \mathcal{C}} \omega_{i,j}(t),$$

ce qui revient à rechercher la zone du graphe possédant la **densité** la plus élevée en liaisons fortes. Cette démarche peut être utile, par exemple, lorsque l'on souhaite mettre en évidence un noyau collaboratif de grande intensité.

Un **deuxième** critère, plus nuancé, repose sur le **rapport** entre la force interne et la force externe. On introduit alors une fonction de ratio

$$R(\mathcal{C}) = \frac{\sum_{i,j \in \mathcal{C}} \omega_{i,j}(t)}{\sum_{i \in \mathcal{C}, j \notin \mathcal{C}} \omega_{i,j}(t) + \epsilon},$$

où $\epsilon > 0$ agit comme un **terme de régularisation** évitant les divisions par zéro. Le cluster \mathcal{C} optimal maximise alors $R(\mathcal{C})$, ce qui revient à privilégier des groupes dont la **cohésion interne** se révèle importante face aux connexions dirigées vers l'extérieur.

Enfin, un **troisième** point de vue fait appel à une **fonction d'énergie** \mathcal{J} , ou fonction de coût, qui **récompense** la densité intra-cluster tout en **pénalisant** les connexions externes (cf. section 1.4.2 pour une présentation plus générale du concept de coût dans le DSL). Les **clusters** apparaissent alors comme des **minima locaux** de \mathcal{J} . Cette perspective offre un cadre théorique aisément connectable à la physique statistique ou aux techniques d'optimisation combinatoire, permettant d'étudier la **stabilité** ou le **caractère** global/local des solutions.

Malgré la diversité de ces approches, le principe fondamental reste le même. Un **cluster** est un **groupe d'entités** dont les liaisons, au sein de la matrice $W(t)$, sont relativement plus fortes entre ses membres qu'avec le reste du réseau. On retrouve ainsi l'idée intuitive d'une **collaboration** accrue à l'intérieur du sous-ensemble, associée à un **isolement** ou à une **différenciation** face aux entités extérieures. Dans la suite du texte, ces définitions de clusters s'avèrent essentielles pour analyser la manière dont le **SCN** laisse émerger automatiquement, au fil de la dynamique d'apprentissage, des **strates** coopératives plus ou moins vastes.

Dans le cadre du **Deep Synergy Learning**, les pondérations $\omega_{i,j}(t)$ évoluent fréquemment selon la règle

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)],$$

où $\eta > 0$ et $\tau > 0$ sont des constantes, tandis que $S(\mathcal{E}_i, \mathcal{E}_j)$ désigne la **synergie** entre les entités i et j . Dès lors qu'une pondération $\omega_{i,j}(t)$ dépasse un seuil ω_{\min} , on considère qu'il existe un **lien effectif** entre les deux entités. Plusieurs phénomènes concourent alors à la **formation d'un cluster** :

- **Renforcement interne.** Lorsqu'un certain **groupe** $\mathcal{C} \subset \{1, \dots, n\}$ présente des synergies élevées entre ses membres, les pondérations $\omega_{i,j}(t)$ associées ont tendance à **augmenter** pour tous $i, j \in \mathcal{C}$.
- **Isolement progressif.** À mesure que les liens internes se consolident, les connexions externes (i.e. celles dont la synergie avec les entités extérieures demeure plus faible) s'**atténuent**, rendant le sous-ensemble \mathcal{C} de plus en plus autonome et homogène.

Ce processus engendre un **cluster** cohérent, formé sans qu'aucun agencement préalable ne soit imposé.

Exemple de scénario d'émergence

Temps initial (t=0). Les pondérations $\omega_{i,j}(0)$ sont faibles ou distribuées de manière uniforme, si bien qu'aucune structure évidente ne se détache.

Interactions locales. Certaines entités détectent une synergie ($S(\mathcal{E}_i, \mathcal{E}_j) > 0$) suffisamment avantageuse pour se renforcer localement. Les $\omega_{i,j}(t)$ correspondantes s'amplifient.

Formation d'une "graine". Progressivement, un **mini-cluster** naît sous la forme de 2 ou 3 entités dont les liens internes, devenus plus élevés, assurent un noyau de haute synergie.

Attraction. Si une entité \mathcal{E}_k à la périphérie présente aussi une bonne synergie avec cette graine, ses liaisons $\omega_{k,i}$ ($i \in \mathcal{C}$) se renforcent, ce qui intègre \mathcal{E}_k au cluster en expansion.

Consolidation. Une fois que les **liens internes** atteignent un niveau stable et que les connexions vers l'extérieur restent faibles, le groupe \mathcal{C} se **stabilise** comme cluster. D'un point de vue formel, cela correspond à un **minimum local** d'une fonction d'énergie \mathcal{J} , où \mathcal{C} se comporte comme un **attracteur**.

Au-delà des **clusters** de base, il n'est pas rare que plusieurs sous-groupes finis se **rassemblent** pour former un **macro-cluster**, c'est-à-dire une entité supérieure englobant plusieurs clusters déjà constitués.

Pour cela, on considère un **méta-groupe** \mathcal{M} rassemblant des **clusters** $\mathcal{C}_1, \mathcal{C}_2, \dots$. Les pondérations entre deux clusters \mathcal{C}_1 et \mathcal{C}_2 se définissent par

$$\Omega(\mathcal{C}_1, \mathcal{C}_2) = \frac{1}{|\mathcal{C}_1| \cdot |\mathcal{C}_2|} \sum_{\substack{i \in \mathcal{C}_1 \\ j \in \mathcal{C}_2}} \omega_{i,j}(t).$$

Si $\Omega(\mathcal{C}_1, \mathcal{C}_2)$ dépasse un certain niveau, les deux groupes **fusionnent**, produisant ainsi un **macro-cluster** élargi.

Le **processus de fusion** des clusters suit une dynamique progressive où les liaisons synergiques évoluent jusqu'à structurer des **macro-clusters** plus larges.

Les **clusters initiaux**, notés \mathcal{C}_1 et \mathcal{C}_2 , existent indépendamment dans le réseau. Lorsque plusieurs connexions $\omega_{i,j}$, avec $i \in \mathcal{C}_1$ et $j \in \mathcal{C}_2$, se renforcent progressivement, la synergie globale entre ces deux groupes augmente. Ce **renforcement inter-clusters** favorise leur rapprochement et amplifie leur interdépendance.

Dès que ce renforcement atteint un **seuil critique**, la distinction entre \mathcal{C}_1 et \mathcal{C}_2 disparaît. Ils cessent alors d'être des sous-groupes autonomes et forment un **macro-cluster** \mathcal{M} , intégrant leurs dynamiques respectives.

Ce mécanisme s'apparente au **clustering hiérarchique agglomératif**, à la différence que dans le **DSL**, les pondérations $\omega_{i,j}$ ne sont pas statiques. Elles s'ajustent en permanence, permettant une **consolidation dynamique** qui s'adapte aux évolutions du réseau et aux changements des interactions synergiques.

Pour formaliser ces phénomènes de clusterisation, on introduit couramment une **fonction d'énergie** \mathcal{J} :

$$\mathcal{J}(\{\omega_{i,j}\}) = - \sum_{(i,j)} \omega_{i,j} S(\mathcal{E}_i, \mathcal{E}_j) + \alpha \sum_{(i,j)} (\omega_{i,j})^2 + \dots$$

Le premier terme **récompense** les liaisons présentant une **haute synergie**. Minimiser \mathcal{J} revient ainsi à **maximiser** la somme $\sum \omega_{i,j} S(\mathcal{E}_i, \mathcal{E}_j)$, favorisant le renforcement des connexions les plus pertinentes.

Le second terme, avec $\alpha > 0$, introduit une **régularisation** ou un **terme de pénalité**, limitant une croissance excessive et simultanée de toutes les connexions. Cette contrainte prévient une surdensification du réseau et maintient un équilibre dans la distribution des pondérations.

Dans ce cadre, un **cluster** renvoie à un **minimum local** de \mathcal{J} , où l'on observe un niveau élevé de pondérations **en interne**. Un **macro-cluster** surgit quand on identifie un autre minimum local, rassemblant plusieurs groupes préexistants.

L'émergence de **clusters** segmente le Synergistic Connection Network en **régions** spécialisées (p. ex. un cluster associant vision et texte, un autre combinant audio et capteurs, etc.). On obtient ainsi une forme de **division du travail** au sein d'un réseau de grande taille, ce qui favorise la **lisibilité** et la **résilience** globale.

Par ailleurs, la formation de **macro-clusters** peut être rapprochée de l'idée de **modules cognitifs**, dans lesquels plusieurs sous-ensembles d'entités se lient pour s'attaquer à des tâches plus complexes. Le réseau acquiert de la sorte une **modularité adaptative**, qu'il s'agisse de manipuler différents canaux de données ou d'élaborer des représentations de plus haut niveau.

D'un point de vue **algorithmique**, l'identification et la gestion explicite des **clusters** permettent d'optimiser la **mise à jour** des pondérations. Plutôt que d'analyser l'ensemble du graphe simultanément, il devient possible de limiter certains calculs aux **modules faiblement connectés**, réduisant ainsi la complexité computationnelle tout en préservant la dynamique adaptative du réseau. Cette **auto-organisation** en groupes soudés se démarque d'un simple "clustering statique", car les pondérations $\omega_{i,j}(t)$ évoluent continuellement, autorisant des scissions, des fusions et même la disparition de clusters entiers au gré des fluctuations de la synergie.

Exemples d'application

Les **applications du SCN** exploitent sa capacité d'auto-organisation et d'adaptation dynamique pour structurer les interactions entre entités en fonction de la synergie détectée.

Dans l'**analyse multimodale en temps réel**, un système de surveillance peut regrouper en **clusters dynamiques** les entités traitant l'image et la parole, formant par exemple un cluster "visage + discours". Un autre cluster peut émerger pour regrouper les entités chargées de détecter les mouvements anormaux ou les sons inhabituels. Ces agrégats évoluent continuellement en fonction de la **synergie perçue** entre différents flux, tels que la lumière, le son et les déplacements.

Dans la **recommandation et le filtrage collaboratif**, les utilisateurs et les contenus sont représentés par des entités connectées par des pondérations $\omega_{u,c}$. Lorsqu'un utilisateur u apprécie un contenu c , leur connexion se renforce, favorisant l'émergence de **clusters** d'utilisateurs aux préférences similaires. Avec le temps, un **macro-cluster** peut apparaître, regroupant un ensemble d'utilisateurs et les contenus correspondant collectivement à leurs intérêts.

Dans le **traitement biologique ou neuroscientifique**, les neurones ou de petites populations neuronales peuvent être représentés comme des entités \mathcal{E}_i . L'auto-organisation du SCN conduit alors à la formation de **clusters neuronaux locaux**, puis à l'émergence de **macro-clusters** reliant plusieurs aires cérébrales interconnectées. Ce phénomène reflète une **organisation fonctionnelle** à plus grande échelle, traduisant la dynamique sous-jacente des interactions neuronales.

1.4.4. Fonctions de Synergie : Distance, Similarité et Co-Information

Au cœur du **Deep Synergy Learning (DSL)**, les *pondérations synergiques* $\omega_{i,j}$ reliant deux entités \mathcal{E}_i et \mathcal{E}_j évoluent en fonction d'une **mesure de synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$. Cette fonction S évalue dans quelle mesure la **coopération** entre deux entités apporte un **gain** supérieur à l'utilisation isolée de leurs informations.

Bien que le **principe** de la synergie demeure le même (capturer l'apport mutuel), il existe **plusieurs manières** de la **définir** et de la **calculer** dans la pratique. Les plus communes reposent sur (1) des **distances** entre représentations, (2) des **similarités** (souvent normalisées), ou (3) des mesures d'**information** (entropie, co-information, etc.). Cette section expose en détail ces approches, avec une formulation mathématique approfondie.

Dans le **Deep Synergy Learning (DSL)**, la pondération $\omega_{i,j}(t)$ reliant deux entités \mathcal{E}_i et \mathcal{E}_j suit fréquemment une équation de la forme

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)],$$

où η et τ sont des hyperparamètres positifs, et $S(\mathcal{E}_i, \mathcal{E}_j)$ la **fonction de synergie**. Cette fonction reflète dans quelle mesure les entités \mathcal{E}_i et \mathcal{E}_j estiment pouvoir coopérer avantageusement. Dès lors,

- Si $S(\mathcal{E}_i, \mathcal{E}_j)$ est **positif** et assez grand, la liaison $\omega_{i,j}$ se **renforce**.
- Si $S(\mathcal{E}_i, \mathcal{E}_j)$ est faible ou négatif, la liaison s'**affaiblit**.

Le **choix** de la fonction S influe directement sur la **topologie** finale du réseau. Deux entités jugées proches ou complémentaires se retrouveront dans un même **cluster** (voir section 1.4.3), tandis que des entités peu compatibles verront leur lien s'étioler. Différentes approches existent pour définir cette synergie.

1. Synergie fondée sur la distance

Une façon intuitive d'instaurer la synergie est de partir d'une **distance** $d(\mathcal{E}_i, \mathcal{E}_j)$ calculée sur leurs représentations, souvent des vecteurs $\mathbf{x}_i, \mathbf{x}_j \in \mathbb{R}^d$. Un exemple répandu consiste en la **distance euclidienne** $\|\mathbf{x}_i - \mathbf{x}_j\|$. On convertit ensuite cette distance en un score de synergie **décroissant** :

$$S(\mathcal{E}_i, \mathcal{E}_j) = \frac{1}{1 + \|\mathbf{x}_i - \mathbf{x}_j\|^2},$$

ou bien, en autorisant des valeurs négatives, on peut poser

$$S(\mathcal{E}_i, \mathcal{E}_j) = -\|\mathbf{x}_i - \mathbf{x}_j\|.$$

L'idée est qu'une faible distance se traduit par une **forte synergie**, tandis qu'une distance importante produit un score négatif ou tendant vers zéro. Il est fréquent d'ajouter un **paramètre d'échelle** $\sigma > 0$ pour réguler la sensibilité aux écarts, par exemple

$$S(\mathcal{E}_i, \mathcal{E}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / 2\sigma^2).$$

Cette fonction valorise la synergie pour les entités **très proches** et la fait décroître rapidement au-delà d'un rayon σ . La distance choisie dépend de la nature des données. Elle peut être de type **L1**, **L2**, ou encore définie sur des **tenseurs d'images** ou des **spectrogrammes**, selon le contexte d'application.

2. Synergie fondée sur la similarité

Un autre choix repose sur une **similarité** plutôt que sur une distance. Par exemple, la **similarité cosinus** :

$$\text{sim}(\mathbf{x}_i, \mathbf{x}_j) = \frac{\langle \mathbf{x}_i, \mathbf{x}_j \rangle}{\|\mathbf{x}_i\| \|\mathbf{x}_j\|},$$

qui varie dans $[-1, 1]$. On peut alors définir

$$S(\mathcal{E}_i, \mathcal{E}_j) = \max(0, \text{sim}(\mathbf{x}_i, \mathbf{x}_j)) \quad \text{ou} \quad S(\mathcal{E}_i, \mathcal{E}_j) = 1/2 (\text{sim}(\mathbf{x}_i, \mathbf{x}_j) + 1),$$

histoire de normaliser la valeur dans $[0, 1]$. Cette approche est pratique quand on souhaite regrouper des **vecteurs** se ressemblant en direction ou quand on manipule des **corrélations** (ex. coefficient de Pearson). Elle est parfois moins adaptée si l'on cherche à capter la complémentarité non linéaire entre des entités fortement dissemblables.

3. Synergie fondée sur la co-information ou l'information mutuelle

Lorsque les entités $\mathcal{E}_i, \mathcal{E}_j$ renvoient à des **variables aléatoires** $\mathbf{X}_i, \mathbf{X}_j$, on peut évaluer la synergie via des **quantités entropiques** :

$$I(\mathbf{X}_i; \mathbf{X}_j) = H(\mathbf{X}_i) + H(\mathbf{X}_j) - H(\mathbf{X}_i, \mathbf{X}_j),$$

qui mesure l'**information mutuelle**. Toutefois, cette mesure ne différencie pas la simple **redondance** d'une véritable **coopération**. On peut alors recourir à la **co-information** ou à la **Partial Information Decomposition (PID)**, laquelle décompose l'information partagée en portions "synergique" et "redondante". Cela donne un score :

$$S(\mathcal{E}_i, \mathcal{E}_j) = \max(0, I_{\text{syn}}(\mathbf{X}_i, \mathbf{X}_j \mid \mathbf{Y})),$$

où I_{syn} désigne la contribution strictement synergique, au sens de la définition employée (co-info, PID, etc.). Cette méthode peut détecter des interactions non linéaires et complexes, mais elle se révèle plus **onéreuse** à calculer, en particulier pour des données de haute dimension, et nécessite de **choisir** la mesure d'information la plus pertinente au contexte.

Conseils pour le choix de la fonction S

Le **DSL** n'impose pas une unique forme de synergie, mais propose un **cadre flexible** où toute mesure d'**apport mutuel** entre entités peut être adaptée. Certains principes guident néanmoins le choix de la fonction S .

La **distance** est une option simple à implémenter et intuitive. Elle favorise l'association d'entités similaires dans un certain espace de représentation, mais peut sous-estimer la complémentarité

entre des entités très différentes mais hautement coopératives. La **similarité** est plus pertinente lorsque l'on compare des directions ou des vecteurs normalisés. Une forte similarité indique une parenté marquée, bien que certaines complémentarités non linéaires puissent lui échapper.

L'**information mutuelle**, la **co-information** ou encore le **PID** permettent de capturer des **coopérations complexes et non linéaires**. Ces approches sont puissantes mais présentent un coût computationnel élevé et restent sensibles aux méthodes d'estimation statistique.

Il est évidemment possible de **combiner** plusieurs indicateurs (distance, similarité, information mutuelle) ou de pondérer différentes composantes pour confectionner un score $S(\cdot, \cdot)$ plus adapté. Le choix exact dépendra de la nature des données, du domaine d'application et des **objectifs** poursuivis (clustering, prédiction, fusion multimodale, etc.).

Exemples d'utilisation

Analyse d'images. Considérons des entités \mathcal{E}_i représentant des **patches** ou des **régions d'images**. Une distance **L2** appliquée aux vecteurs de caractéristiques, extraits par exemple d'un **réseau de neurones convolutionnel**, peut fournir un score de **synergie décroissant**. Plus les patches présentent une forte ressemblance, plus leur coopération est renforcée.

Fusion multimodale. Supposons deux entités $\mathbf{X}_{\text{audio}}$ et $\mathbf{X}_{\text{visuel}}$. On mesure leur co-information conditionnellement à une variable \mathbf{Y} (classe d'événement). Si leur combinaison apporte un gain clair, on augmente la synergie et, par conséquent, on renforce leur lien.

Système hybride symbolique-connexionniste. Un module symbolique (décrivant des règles ou faits abstraits) et un module vectoriel (des embeddings d'images, par exemple) peuvent relier leurs entités si l'information mutuelle ou la similarité cosinus démontre qu'ils s'améliorent réciproquement.

Conclusion et ouverture

La fonction de synergie $S(\mathcal{E}_i, \mathcal{E}_j)$ est la **pièce angulaire** du Deep Synergy Learning. Selon qu'on s'appuie sur une **distance**, une **similarité** ou une **information mutuelle**, on oriente la manière dont les entités **coopèrent** et se **relient** au sein du **Synergistic Connection Network**.

- Les **fonctions de distance** favorisent l'agrégation d'éléments proches dans l'espace des caractéristiques.
- Les **fonctions de similarité** rassemblent des vecteurs corrélés, tout en ignorant parfois la complémentarité.
- Les **mesures entropiques** (co-info, PID, etc.) explorent la coopération profonde, prenant en compte les facettes non linéaires et la contribution uniquement révélée par l'association des entités.

Les sections suivantes (1.4.5 à 1.4.7) approfondiront la **mise à jour temporelle** de $\omega_{i,j}(t)$ fondée sur cette synergie, la distinction entre **interactions directes** et **indirectes**, et la manière de gérer la **synergie n-aire**. L'ensemble de ces points complète la **vision** du DSL comme un réseau évolutif, autonome et apte à faire émerger des **structures** d'apprentissage complexe dans des contextes variés.

1.4.5. Pondérations Adaptatives et Évolution Temporelle

Les sections précédentes (1.4.1 à 1.4.4) ont posé les **fondements** de l'architecture générale du **Deep Synergy Learning (DSL)**. Chaque **entité** peut établir des connexions avec d'autres à travers des **pondérations synergiques** $\omega_{i,j}$, qui traduisent la **valeur ajoutée** de leur coopération. Reste à décrire **comment** ces pondérations **changent** dans le temps, c'est-à-dire la **loi d'évolution** qui fait que le réseau se **reconfigure** en permanence. C'est l'objet de cette sous-section (1.4.5), qui explique :

- Les **équations** gouvernant l'adaptation des pondérations,
- Le sens de cette **dynamique** (renforcement / affaiblissement),
- Les implications sur la **convergence** ou la **stabilité** des liens,
- Le rôle des **paramètres** (taux d'apprentissage, régularisation, etc.) dans l'auto-organisation du réseau.

Cette **évolution temporelle** est au cœur du **SCN** (Synergistic Connection Network), car c'est elle qui autorise la création progressive de **clusters** et, par extension, la modélisation d'un **apprentissage** réellement continu.

Dans la plupart des formulations du **Deep Synergy Learning (DSL)**, la pondération $\omega_{i,j}(t)$ reliant deux entités \mathcal{E}_i et \mathcal{E}_j se met à jour suivant un schéma adaptatif de la forme

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)],$$

où η désigne un **taux d'apprentissage** (ou **pas de mise à jour**) et τ un **coefficient de régulation** (ou *terme d'oubli*). La fonction $S(\mathcal{E}_i, \mathcal{E}_j)$ correspond à la mesure de **synergie** choisie (distance, similarité, co-information, etc.). Ce cadre de mise à jour s'interprète de la manière suivante.

Lorsque la synergie S entre \mathcal{E}_i et \mathcal{E}_j est suffisamment élevée et que $\omega_{i,j}(t)$ reste encore modérée, le terme $\eta [S - \tau \omega]$ demeure **positif**, entraînant un **renforcement** progressif de la liaison à chaque itération.

À l'inverse, si la synergie est faible ou négative, ou si $\omega_{i,j}(t)$ atteint une valeur déjà trop élevée, ce terme devient **négatif**, ce qui conduit à un **affaiblissement** progressif de la liaison.

On peut mettre en évidence un **point fixe** en imposant $\omega_{i,j}(t+1) = \omega_{i,j}(t) = \omega_{i,j}^*$. La condition associée,

$$\omega_{i,j}^* = S/\tau,$$

montre que plus la **synergie** est grande, plus la **pondération d'équilibre** $\omega_{i,j}^*$ atteint un niveau élevé.

Variantes et enrichissements

Plusieurs adaptations sont possibles autour de cette règle. On peut retenir un **historique** afin de lisser la dynamique ou pondérer les valeurs antérieures $\omega_{i,j}(t-1)$. La mise à jour des pondérations peut adopter une **forme non linéaire**, s'écartant du modèle linéaire $\omega + \eta [S - \tau \omega]$ pour inclure des équations **saturantes**, **exponentielles** ou inspirées des **règles hebbiennes**.

De plus, la synergie S peut évoluer dans le temps si les entités $\mathcal{E}_i(t)$ et $\mathcal{E}_j(t)$ mettent à jour leurs représentations internes. Dans ce cas, la pondération $\omega_{i,j}(t)$ n'influence pas seulement la connexion entre entités, mais modifie également la **synergie perçue**, induisant une transformation dynamique de la **structure globale** du réseau.

Système dynamique discret ou continu

Le modèle

$$\omega_{i,j}(t+1) - \omega_{i,j}(t) = \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)]$$

est une **équation aux différences**, interprétable comme un **système dynamique** discret dans un espace de dimension $\binom{n}{2}$ (ou $n(n-1)$ selon que l'on distingue la direction $i \rightarrow j$). On y analyse notamment :

- Les **points fixes** $\omega_{i,j}^*$ satisfaisant

$$\omega_{i,j}^* = \omega_{i,j}^* + \eta [S - \tau \omega_{i,j}^*].$$

- La **stabilité** de ces points. Par exemple, en linéarisant autour de l'équilibre, on obtient

$$\Delta \omega_{i,j}(t+1) = \Delta \omega_{i,j}(t) - \eta \tau \Delta \omega_{i,j}(t),$$

montrant que la condition $\eta \tau < 1$ est usuellement requise pour maintenir une évolution stable des pondérations.

L'équivalent **en temps continu** est donné par l'ODE

$$d \omega_{i,j} / dt = \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)].$$

Ici encore, la solution tend vers $\omega_{i,j}^* = S/\tau$ s'il ne survient pas d'autres perturbations ni dépendances supplémentaires.

Seuils, cap et parsimonie

Dans le **DSL**, on introduit souvent un **seuil** ω_{\min} en deçà duquel la connexion est jugée inexistante :

$$\omega_{i,j}(t) < \omega_{\min} \Rightarrow \text{aucun lien effectif entre } i \text{ et } j.$$

Même si l'équilibre $\omega_{i,j}^*$ est légèrement positif, la liaison doit dépasser ω_{\min} pour être considérée. À l'autre extrémité, on peut imposer un **cap** $\omega_{i,j}(t) \leq \omega_{\max}$ pour freiner l'augmentation des pondérations et éviter les valeurs extrêmes. Sur un **diagramme de phase** ($\omega - \dot{\omega}$), ces deux seuils

assurent une **parcimonie** structurale. Seuls les liens ayant démontré une **synergie conséquente** sont préservés, tandis que ceux restant trop faibles ou atteignant des valeurs excessives sont soit **coupés**, soit **plafonnés** pour éviter une sur-saturation du réseau.

Interdépendance globale

Dans un réseau comptant n entités, la synergie $S(\mathcal{E}_i, \mathcal{E}_j)$ peut dépendre d'autres entités (cf. section 1.4.4). On écrit parfois

$$S(\mathcal{E}_i, \mathcal{E}_j \mid \{\mathcal{E}_k\}_{k \neq i,j}),$$

pour souligner le fait qu'une **co-information** conditionnelle peut influencer sur la pondération. Les **misés à jour** des pondérations $\omega_{i,j}$ sont **couplées**, ce qui signifie que la liaison $i \leftrightarrow j$ évolue en interaction avec d'autres connexions comme $j \leftrightarrow k$, influençant ainsi la dynamique globale du réseau. Ce couplage favorise l'**émergence de clusters** (section 1.4.3). Les entités d'un sous-groupe \mathcal{C} stabilisent progressivement leurs pondérations internes, formant une **bulle de coopération élevée**, tandis que les connexions extérieures, moins profitables, s'atténuent naturellement.

Feedback et plasticité

Interpréter la dynamique du **DSL** comme un **système à feedback** apporte un éclairage pertinent sur son mode d'auto-organisation.

Le **feedback positif** intervient lorsqu'une liaison $\omega_{i,j}$ se renforce sous l'effet d'une synergie positive ($S > 0$). Dans ce cas, \mathcal{E}_i et \mathcal{E}_j augmentent leur coopération, consolidant ainsi encore davantage leur connexion.

À l'inverse, le **feedback négatif** empêche les liaisons de croître indéfiniment. Si une connexion s'emballe, le terme $-\tau \omega_{i,j}(t)$ ou une autre forme de régulation, comme une pénalisation non linéaire, freine son amplification, garantissant ainsi la **stabilisation** du réseau.

De façon analogique, on peut s'inspirer de la **plasticité synaptique** :

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta \left(a_i(t) a_j(t) \right) - \alpha \omega_{i,j}(t),$$

où $a_i(t)$ et $a_j(t)$ sont les "activités" ou "états" des entités, vus comme un substitut instantané à la synergie $S_{i,j}(t)$. Par ailleurs, l'usage d'un **terme cubique** $-\beta (\omega_{i,j})^3$ agit comme une stabilisation encore plus marquée, évitant la croissance sans borne.

Impacts et conclusions

Flexibilité continue. Grâce à l'actualisation en continu de $\omega_{i,j}(t)$, le **Synergistic Connection Network** (SCN) s'adapte sans cesse aux nouveaux flux ou aux modifications d'environnement, bien plus qu'un réseau à architecture fixe.

- **Formation de clusters.** Les entités liées par une synergie forte consolident leurs liaisons, s'assemblent en **clusters**, et peuvent se **fuser** en macro-clusters (voir section 1.4.3). Cette recomposition est fluide et dépend directement de la dynamique des pondérations.

- **Régime stable vs. oscillatoire.** Selon les choix de synergie, de η , de τ , etc., le réseau peut tendre vers un état stable, osciller, ou coexister dans plusieurs configurations d'équilibre.
- **Aspects computationnels.** Mettre à jour toutes les paires (i, j) exige un coût a priori en $O(n^2)$. Des heuristiques (sparse updates, random sampling) peuvent s'avérer nécessaires pour gérer de grands n de façon **scalable**.

Cette **dynamique adaptative** des pondérations constitue le moteur de l'**auto-organisation** dans le DSL, permettant de découvrir et de consolider des **sous-structures** au sein du réseau. Les sections suivantes (1.4.6 et 1.4.7) s'intéresseront à la **distinction entre interactions directes et indirectes** et à la **synergie n-aire**, deux aspects cruciaux pour comprendre la **richesse** et la **réactivité** de ce paradigme d'apprentissage.

1.4.6. Interactions Directes et Indirectes

Jusqu'à présent, la description du **Deep Synergy Learning (DSL)** s'est essentiellement focalisée sur la **relation directe** entre deux entités \mathcal{E}_i et \mathcal{E}_j , mesurée par une **synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$ et traduite dans une **pondération** $\omega_{i,j}$. Dans un réseau où de nombreuses entités coexistent, les **interactions indirectes** sont essentielles. Même sans liaison directe, deux entités peuvent s'influencer si elles sont **connectées** par une **chaîne** de liens synergiques. Ce mécanisme favorise la propagation des influences et la structuration dynamique du réseau. Cette section (1.4.6) explique :

- Comment les **interactions** peuvent se propager à travers le **Synergistic Connection Network (SCN)**,
- Comment la **synergie** entre deux entités peut être **modulée** par leurs relations avec d'autres,
- Quelles implications mathématiques découlent de ces boucles d'influence plus complexes, notamment pour la **formation** de clusters et l'**émergence** de comportements globaux.

Cette problématique est essentielle pour comprendre l'**auto-organisation**. Même si chaque entité n'agit qu'en fonction de ses liens directs, l'ensemble du réseau peut révéler des effets d'**influence à distance** ou de **coordination** en chaîne, façonnant ainsi sa dynamique globale.

On appelle **interaction indirecte** entre \mathcal{E}_i et \mathcal{E}_j toute séquence $\mathcal{E}_i \rightarrow \mathcal{E}_k \rightarrow \dots \rightarrow \mathcal{E}_m \rightarrow \mathcal{E}_j$ dans laquelle chacune des paires $(\mathcal{E}_u, \mathcal{E}_v)$ du chemin possède une pondération $\omega_{u,v}(t)$ significative. Le **Synergistic Connection Network (SCN)** autorise ainsi la propagation d'**influence** ou de **coopération** le long d'un **chemin** reliant deux entités, même si ces dernières ne sont pas directement connectées.

1. Principes généraux : chemins et relai d'information

Lorsque la liaison $\omega_{k,m}$ est élevée, l'entité \mathcal{E}_k peut **transmettre** (ou "relayer") certaines données ou incitations à \mathcal{E}_m . C'est ainsi qu'une synergie faible ou nulle entre \mathcal{E}_i et \mathcal{E}_j peut progressivement **s'intensifier** si elles interagissent via des entités communes $\{\mathcal{E}_k, \mathcal{E}_\ell, \dots\}$ qui procurent un **gain mutuel**. De même, si l'on modélise la **synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$ comme une quantité **contextuelle**, on peut définir

$$S(\mathcal{E}_i, \mathcal{E}_j \mid \mathcal{E}_k),$$

qui tient compte du fait que la coopération entre \mathcal{E}_i et \mathcal{E}_j peut être **facilitée** ou **rendue moins nécessaire** par la présence d'une entité pivot \mathcal{E}_k .

Exemple : pivot et redondance

- Si \mathcal{E}_k remplit déjà les rôles combinés de \mathcal{E}_i et \mathcal{E}_j , l'intérêt pour \mathcal{E}_i et \mathcal{E}_j de se lier **directement** diminue (redondance).
- Au contraire, si \mathcal{E}_k fournit un complément utile à \mathcal{E}_i et \mathcal{E}_j , ce **contexte** peut augmenter leur synergie bilatérale, incitant l'émergence d'un lien direct $\omega_{i,j}$.

2. Notion de chemins et puissance de la matrice W

Le **graphe** du SCN se décrit par la matrice pondérée $W(t)$, de taille $n \times n$, dont l'entrée $\omega_{i,j}(t)$ indique la connexion directe entre \mathcal{E}_i et \mathcal{E}_j . Lorsqu'on considère un **chemin** de longueur 2 entre \mathcal{E}_i et \mathcal{E}_j via un intermédiaire \mathcal{E}_k , la multiplication de matrices révèle que $(W^2)_{i,j} = \sum_{k=1}^n \omega_{i,k}(t) \omega_{k,j}(t)$. De manière analogue, $(W^p)_{i,j}$ réunit l'effet de **tous** les chemins de longueur p liant i à j .

Si la dynamique des pondérations reste **relativement stable**, l'examen des puissances W^2, W^3, \dots peut dévoiler des **interactions indirectes** (chemins multiples ou plus longs). Des valeurs élevées de $(W^p)_{i,j}$ traduisent une **forte influence** de \mathcal{E}_i sur \mathcal{E}_j à travers plusieurs intermédiaires. En présence de **cycles** (par exemple, $\mathcal{E}_i \rightarrow \mathcal{E}_k \rightarrow \mathcal{E}_m \rightarrow \mathcal{E}_i$), des effets d'**amplification** peuvent apparaître. Si chaque liaison du cycle se renforce, la synergie circule en boucle, entraînant une augmentation progressive de son intensité au sein du réseau.

D'un point de vue **linéaire**, la **valeur propre** de plus grande amplitude, correspondant au **rayon spectral** $\rho(W)$, détermine la **stabilité** du système.

Si $\rho(W) > 1$, de petites perturbations internes peuvent se propager et s'amplifier, entraînant un risque d'**oscillations** ou d'emballement du réseau.

Si $\rho(W) < 1$ après régulation, la dynamique tend à converger vers un **état stable**, maintenant ainsi l'équilibre du système.

3. Contexte et synergie conditionnelle

Dans un réseau synergique, l'existence de **chemins indirects** modifie considérablement la manière dont deux entités $\mathcal{E}_i, \mathcal{E}_j$ perçoivent leur **coopération**. Même si $\omega_{i,j}$ démarre faible, il peut se consolider via l'influence d'entités tierces \mathcal{E}_k . Formellement, on modélise cette dépendance par une fonction

$$S(\mathcal{E}_i, \mathcal{E}_j \mid \{\omega_{i,k}, \omega_{k,j}, \dots\}),$$

introduisant un **système dynamique non linéaire**. Les équations des pondérations deviennent

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta [S(\mathcal{E}_i, \mathcal{E}_j | \{\omega_{i,k}, \omega_{k,j}, \dots\}) - \tau \omega_{i,j}(t)].$$

La dépendance entre les $\omega_{i,j}$ fait naître des **boucles de feedback**, où un renforcement dans un segment du réseau rejaillit sur d'autres liaisons, et ainsi de suite. En pratique, l'**analyse** de stabilité ou de convergence peut impliquer des techniques avancées (théorie des bifurcations, méthodes de Lyapunov, etc.).

4. Rôle des interactions indirectes dans la dynamique globale

Les **interactions indirectes** permettent à un réseau **DSL** d'évoluer de manière dynamique à travers plusieurs mécanismes.

Elles facilitent l'**émergence de nouveaux liens**. Si deux entités \mathcal{E}_i et \mathcal{E}_j constatent qu'elles coopèrent efficacement via un pivot \mathcal{E}_k , elles peuvent établir une connexion directe $\omega_{i,j} > 0$, optimisant ainsi leurs échanges.

Elles favorisent la **fusion des clusters**. Deux groupes initialement distincts, reliés par des chemins transitifs, peuvent progressivement reconnaître leur complémentarité et **coalescer** en un **macro-cluster**, renforçant la structure globale du réseau.

Elles influencent également la **gestion du bruit**. Une entité bruyante ou imprécise peut affecter des voisins même en l'absence de connexion directe, tandis qu'un intermédiaire de confiance peut stabiliser et améliorer la fiabilité des échanges entre plusieurs entités.

Ce mécanisme d'**influence à distance** se retrouve dans des domaines variés. En **robotique multi-agents**, un agent \mathcal{E}_i peut influencer un autre \mathcal{E}_j via un agent \mathcal{E}_k jouant le rôle de relai ; la structure d'échange se reconfigure selon la performance observée. En **neurosciences**, de nouvelles synapses directes se forment après une longue période où l'activité neuronale transite par plusieurs synapses intermédiaires.

Conclusion et ouverture

Les **interactions indirectes** forment un pivot crucial pour saisir l'**auto-organisation** d'un Synergistic Connection Network (SCN). Deux entités \mathcal{E}_i et \mathcal{E}_j initialement faiblement connectées peuvent se **rapprocher** grâce aux chemins multiples les reliant via des tiers, tout comme des boucles internes peuvent amplifier ou réguler la synergie. Cette dimension de **propagation** et de **contexte** est essentielle pour comprendre la **formation** de macro-structures, la dynamique des **cycles** et la **coévolution** des pondérations dans le DSL.

La section suivante (1.4.7) examinera de plus près la **synergie binaire** et surtout la **synergie n-aire**, c'est-à-dire l'impact de coopérations impliquant simultanément plus de deux entités, un aspect qui renforce encore la **richesse** des comportements émergents dans le Deep Synergy Learning.

1.4.7. Synergie binaire et n-aire : au-delà des relations deux à deux

La **synergie binaire** demeure la forme la plus couramment utilisée pour évaluer le degré de coopération entre deux entités \mathcal{E}_i et \mathcal{E}_j . On définit pour cela une fonction $S(\mathcal{E}_i, \mathcal{E}_j)$ (pouvant être fondée sur une distance, une similarité, de la co-information, etc.) qui mesure dans quelle mesure

$$\mathbf{x}_i (+ \mathbf{s}_i, \theta_i) \Rightarrow \mathbf{x}_j (+ \mathbf{s}_j, \theta_j)$$

s'enrichissent mutuellement ou, au contraire, s'avèrent redondantes ou peu utiles l'une à l'autre.

- **Mise à jour.** À chaque itération, la pondération $\omega_{i,j}(t)$ se voit recalculée en fonction de la valeur $S(\mathcal{E}_i, \mathcal{E}_j)$.
- **Clusters.** Les clusters décrits en section 1.4.3 émergent généralement d'une **somme** ou d'une **agrégation** de ces liaisons binaires (souvent assortis d'un seuil, stabilisant la formation de groupes plus étendus).

Toutefois, la synergie binaire suppose que l'apport mutuel entre \mathcal{E}_i et \mathcal{E}_j ne dépende **pas** explicitement des autres entités. Or, la section 1.4.6 a mis en évidence l'influence du **contexte global** sur les interactions du réseau. La présence d'entités intermédiaires comme \mathcal{E}_k, \dots peut modifier significativement les dynamiques locales. Certaines lacunes peuvent être **comblées**, tandis que des informations complémentaires peuvent se **mutualiser**, révélant ainsi une **coopération émergente** impliquant plusieurs entités simultanément.

1. Synergie n-aire : idée d'un effet collectif

La **synergie n-aire** engage un **ensemble** $\{\mathcal{E}_{k_1}, \dots, \mathcal{E}_{k_m}\}$ (avec $m \geq 3$ dans la plupart des cas), dont la **combinaison** fournit une valeur ajoutée qu'on ne saurait réduire à la somme des synergies binaires. On écrit

$$S_n(\mathcal{E}_{k_1}, \dots, \mathcal{E}_{k_m})$$

pour désigner la synergie qui s'exprime **simultanément** entre ces m entités. On veut y saisir l'idée que "toutes ensemble, elles valent davantage que la somme de leurs collaborations deux à deux". Pour le cas de trois entités $\mathcal{E}_a, \mathcal{E}_b, \mathcal{E}_c$, on parle parfois de **complémentarité stricte** :

$$S_n(\mathcal{E}_a, \mathcal{E}_b, \mathcal{E}_c) > S(\mathcal{E}_a, \mathcal{E}_b) + S(\mathcal{E}_b, \mathcal{E}_c) + S(\mathcal{E}_a, \mathcal{E}_c),$$

signifiant que le **trio** dans son intégralité apporte plus que n'importe quelle addition de paires isolées. Pour $m > 3$, ce principe s'étend à une **coopération collective**, où l'interaction ne se limite plus à des relations binaires, mais implique plusieurs entités agissant de concert.

En **théorie de l'information**, on retrouve cette distinction dans :

- **Information mutuelle totale.**

$$I(\mathbf{X}_1, \dots, \mathbf{X}_m) = \sum_{i=1}^m H(\mathbf{X}_i) - H(\mathbf{X}_1, \dots, \mathbf{X}_m),$$

laquelle ne fait pas le tri entre redondance et synergie réellement n-aire.

- **Partial Information Decomposition (PID).**

Ce formalisme sépare la part d'information partagée par tous (redondance) de la part spécifiquement **synergique**, c'est-à-dire l'information qui n'apparaît qu'en combinant l'ensemble complet des variables $\mathbf{X}_1, \dots, \mathbf{X}_m$. S'il émerge une synergie n-aire positive, on découvre un **surcroît** d'information inexistant dans tout sous-groupe plus restreint.

Pour le **DSL**, ces notions indiquent comment un cluster de taille m peut se constituer lorsqu'on détecte une synergie n-aire **positive** (ou dépassant un certain seuil).

2. Intégration de la synergie n-aire dans un réseau

Si les pondérations $\omega_{i,j}$ portent sur des **liaisons binaires**, comment tirer parti d'une synergie n-aire ? Plusieurs pistes :

- **Approche factorisée.**

Créer explicitement une pondération ω_{k_1, \dots, k_m} reliant simultanément les entités $\{\mathcal{E}_{k_1}, \dots, \mathcal{E}_{k_m}\}$. Il faut alors manipuler un **hyper-graphe** (où une "arête" connecte plus de deux nœuds) ou stocker un **tenseur** de pondérations, ce qui devient coûteux si m croît.

- **Approche "bonus/malus".**

Reprendre le cadre binaire $\omega_{i,j}$ en y ajoutant un **terme** lié à la synergie n-aire. Par exemple, pour trois entités $\mathcal{E}_a, \mathcal{E}_b, \mathcal{E}_c$, on écrit :

$$\omega_{a,b}(t+1) = \omega_{a,b}(t) + \eta [S(\mathcal{E}_a, \mathcal{E}_b) + \gamma S_3(\mathcal{E}_a, \mathcal{E}_b, \mathcal{E}_c) - \tau \omega_{a,b}(t)].$$

Ici, γ pèse la contribution de la synergie tripartite. Si $\{\mathcal{E}_a, \mathcal{E}_b, \mathcal{E}_c\}$ forment un trio très coopératif, chaque lien binaire (a-b, b-c, a-c) reçoit un **surcroît** de renforcement, favorisant la cristallisation d'un micro-cluster à trois.

3. Émergence de micro- ou macro-clusters grâce à la synergie n-aire

Lorsqu'un groupe $\mathcal{E} * k_1, \dots, \mathcal{E} * k_m$ présente une **synergie n-aire élevée**, plusieurs dynamiques émergent naturellement.

Les **liens binaires internes** se **renforcent**, consolidant ainsi la connectivité entre les entités du groupe. Cette intensification des interactions conduit à la **stabilisation** d'un sous-graphe quasi complet, traduisant une forte cohésion interne. Enfin, la coopération entre ces entités génère un

gain global qui dépasse la simple somme des contributions individuelles, illustrant la puissance de la synergie collective.

Si, plus tard, un autre sous-groupe $\{\mathcal{E}_{r_1}, \dots, \mathcal{E}_{r_q}\}$ se joint à ce collectif parce que la **synergie** entre ces deux ensembles élargis se révèle à son tour avantageuse, on peut assister à la **fusion** de clusters en un **macro-cluster** plus vaste. L'**approche n-aire** montre que le **DSL** ne se limite pas aux interactions binaires. Il peut détecter qu'un groupe de **quatre, cinq ou dix entités** coopère de manière exceptionnelle, justifiant ainsi l'émergence d'un **module** structuré au sein du réseau.

Cette structure peut être interprétée comme une **hiérarchie** ou un **treillis** organisationnel.

Aux **feuilles**, on retrouve les **synergies binaires** $\omega_{i,j}$, représentant les connexions directes entre deux entités. Plus haut, les **nœuds intermédiaires** correspondent aux **synergies tri-partites, 4-partites**, et ainsi de suite, traduisant des interactions collectives plus complexes. Enfin, aux **sommets**, les **macro-clusters** émergent, englobant de vastes ensembles d'entités interconnectées par des coopérations multiples.

L'**auto-organisation** d'un SCN peut alors être comprise comme un cheminement dans cet espace d'interactions multiples, valorisant les groupes où la **complémentarité** est nettement profitable.

4. Illustrations et implications

Les **illustrations et implications** de cette approche montrent comment la synergie **n-aire** favorise l'émergence de **macro-clusters** et de structures adaptatives.

Dans la **multimodalité**, certaines modalités comme l'audio et le visuel ne présentent pas nécessairement une forte similarité directe. Cependant, l'introduction d'une troisième modalité, comme le texte, peut révéler une **synergie tripartite**. Un **macro-cluster** émerge alors pour gérer simultanément la parole, l'image du locuteur et les informations projetées, optimisant ainsi la compréhension du contenu.

Dans les **systèmes d'agents**, plusieurs robots peuvent former un **équipage coopératif** où l'interaction entre trois ou plusieurs entités permet de résoudre un problème qu'aucune sous-combinaison plus simple ne pouvait traiter efficacement. Le **DSL** capture cette dynamique et facilite la formation de **super-équipes** capables d'adapter leur structure en fonction des besoins.

En **chimie** ou en **biologie**, deux composés A et B peuvent avoir une interaction limitée jusqu'à l'introduction d'un troisième élément C jouant le rôle de **catalyseur**. Ce phénomène révèle une **synergie triple** (A–B–C) qui dépasse toute combinaison restreinte, illustrant l'émergence de propriétés collectives nouvelles.

1.5. Pourquoi une Approche Synergique ?

Les sections précédentes ont mis en évidence la manière dont le **Deep Synergy Learning (DSL)**, à travers son **architecture générale** et son **auto-organisation** (voir sections 1.3 et 1.4), se distingue d'une approche classique. Pour mieux saisir l'intérêt d'une telle approche, il est essentiel de répondre à une question fondamentale. **Pourquoi** adopter un paradigme **synergique** plutôt que de se limiter aux réseaux de neurones profonds ou à d'autres méthodes d'apprentissage ? Cette section 1.5 propose plusieurs raisons majeures :

Avantages par Rapport aux Réseaux Neuronaux Profonds (1.5.1)

Gestion Naturelle de la Multi-modalité (1.5.2)

Flexibilité vis-à-vis des Données Incomplètes ou Bruitées (1.5.3)

Potentiel d'Auto-Évolution et d'Adaptation Continue (1.5.4)

Réduction de la Dépendance à la Supervision Humaine (1.5.5)

Création de Représentations Riches et plus Interprétables (1.5.6)

Intégration de Dimensions Symboliques ou Cognitives (1.5.7)

Chacun de ces points met en lumière les **bénéfices** d'un modèle qui valorise la **co-opération** dynamique entre entités, plutôt qu'une hiérarchie de couches figées. Nous allons tout d'abord (1.5.1) confronter la logique **synergique** du DSL à la structure rigide des **réseaux neuronaux profonds (Deep Learning)**, pour en dégager les avantages potentiels.

1.5.1. Avantages par Rapport aux Réseaux Neuronaux Profonds

Les **réseaux neuronaux profonds** (CNN, RNN, Transformers, etc.) ont assurément démontré leur **efficacité** dans de nombreux domaines. Cependant, le **DSL** propose une **philosophie** radicalement différente, susceptible d'apporter :

1. **Flexibilité Structurelle**
2. **Plastique et Auto-Organisé**
3. **Réduction de la Spécialisation Rigide**
4. **Facilité d'Intégration Multi-entités**
5. **Apprentissage Continu**

Chacun de ces volets constitue un avantage important, que nous détaillons ci-après.

1.5.1.1. Flexibilité Structurale

Dans un **réseau neuronal profond** classique, la **topologie** (nombre et type de couches, neurones, schémas de fusion de données) est entièrement **préétablie** avant l'apprentissage. Par exemple, un **CNN** (Convolutional Neural Network) organise les données selon des couches **convolutionnelles**, suivies éventuellement de **pooling** et de **couches fully-connected**, sans changement d'architecture au cours de l'entraînement. De même, les **Transformers** enchaînent des **blocs self-attention** répétés, toujours selon un canevas décidé en amont. Ces modèles **ne modifient** pas leur structure interne en cours de fonctionnement :

Nombre de couches, arrangement des blocs et synapses demeurent fixes pendant tout l'apprentissage.

Ils **n'autorisent** pas la création ou la **suppression** de neurones ou de liaisons et conservent un agencement identique, même si la **distribution** des données ou le **contexte** subit des modifications significatives.

À l'inverse, dans le **Deep Synergy Learning (DSL)**, on insiste sur la dynamique de ses **entités** et de leurs **liaisons**. Les pondérations $\omega_{i,j}(t)$ entre deux entités \mathcal{E}_i et \mathcal{E}_j évoluent selon la **synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$. Il est possible qu'un **lien** émerge (si la synergie grandit et dépasse un certain **seuil**), qu'il **disparaisse** (si l'intérêt coopératif s'avère trop faible), ou qu'il se **recrée** au cours du temps. La **flexibilité** topologique qui en résulte se formalise par des règles de mise à jour de la forme

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)],$$

où η désigne le **taux d'apprentissage** et τ un **terme de régulation**. Quand la synergie S s'avère positive et assez élevée, la liaison se **renforce** ; dans le cas contraire, elle **s'affaiblit** et peut même s'annuler. Cette dynamique se révèle d'autant plus souple qu'on peut imposer une **parsimonie** via un seuil ω_{\min} ou un **cap** ω_{\max} , comme le rappelle la section 1.4.5.

La conséquence directe de cette **reconfiguration dynamique** tient dans la capacité du réseau à **s'adapter** en continu à l'arrivée de nouvelles données ou à un **changement** de distribution. Au lieu de maintenir invariablement le même chemin de traitement (comme dans un **CNN** figé ou un **Transformers** aux blocs intangibles), le **DSL** autorise une **reliance** directe entre certaines entités apparues pertinentes, tout en relâchant ou en supprimant d'autres connexions devenues inutiles. On peut ainsi écrire, de façon synthétique :

Flexibilité topologique \Rightarrow Capacité d'adaptation et résilience face à l'évolution du contexte.

Dès lors, si la distribution des données subit un changement progressif (par exemple, l'apparition d'une nouvelle modalité sensorielle), le **DSL** peut incorporer de nouvelles **entités** \mathcal{E}_{nov} et renforcer les liaisons $\omega_{\text{nov},j}$ jugées porteuses d'une forte **synergie**. Simultanément, il peut affaiblir ou rompre des liens moins pertinents. C'est précisément cette **malléabilité** qui caractérise le **SCN** et le différencie de la rigidité structurelle des architectures profondes traditionnelles.

Ce **principe** de flexibilité structurelle jette les bases d'une **évolution** continue de la topologie, conforme à l'esprit d'**auto-organisation** présenté dans les sections 1.4.3 et 1.4.5. En ce sens, le **DSL** épouse davantage la **complexité** et la **variabilité** des environnements réels, tout en garantissant que seules les **connexions** profitables (au sens de la synergie) se **consolident** dans le temps.

1.5.1.2. Plasticité et Auto-Organisation

Dans les **réseaux de neurones** usuels, la phase d'apprentissage consiste principalement à **ajuster** les **poids** dans une **architecture** figée à l'avance. En revanche, le **Deep Synergy Learning (DSL)** adopte une approche plus **plastique**. Il permet à la **topologie** du **Synergistic Connection Network (SCN)** de **se transformer** au fil du temps, en ajustant dynamiquement les connexions et les structures en fonction des synergies détectées. Les liaisons $\omega_{i,j}(t)$ qui peinent à démontrer une **synergie** positive ou dont la valeur reste trop faible peuvent **disparaître**, tandis que d'autres connexions, révélant un intérêt mutuel marqué, se **renforcent** jusqu'à forger de nouveaux chemins ou **clusters**.

Pour exprimer cette mise à jour, on recourt le plus souvent à une règle sous la forme

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)],$$

où η représente le **taux d'apprentissage**, τ un terme de **régulation**, et $S(\mathcal{E}_i, \mathcal{E}_j)$ la **synergie** (distance, similarité, information mutuelle...). Ainsi, chaque itération propose une **réévaluation** de la pertinence du lien $\omega_{i,j}$. Si la synergie demeure assez élevée, la liaison persiste ou se consolide ; sinon, elle tend vers zéro et peut être considérée comme **rompue** selon un seuil ω_{\min} (cf. section 1.4.5 pour la gestion des seuils).

Cette **démarche** de régulation évoque la **plasticité synaptique** que l'on observe dans les **systèmes biologiques**. À l'image des neurones qui renforcent leurs connexions en cas de co-activation et en abandonnent d'autres, les entités du DSL remanient leur **graphe** de manière à privilégier les **coopérations** les plus fructueuses. Les **clusters** (sections 1.4.3 et 1.4.4) émergent **spontanément**, sans nécessiter de paramétrage explicite du nombre de groupes. Ils apparaissent dès que les pondérations internes d'un sous-ensemble d'entités atteignent un certain niveau de **synergie**, formant ainsi un **sous-réseau** fortement connecté et structuré.

Cette aptitude à **auto-organiser** la structure s'avère avantageuse pour l'**adaptation** en continu. On peut en effet passer d'une tâche à l'autre ou intégrer une **nouvelle modalité** de données (image, audio, signal textuel) sans reconstruire l'architecture dans son intégralité, puisqu'il suffit d'ajouter des entités et de laisser les liaisons se former ou s'éteindre de façon autonome selon la **synergie** détectée. Les **clusters** demeurant utiles persistent, tandis que de **nouveaux** se créent si des **interactions** inédites se révèlent rentables. Le **réseau** conserve ainsi un **degré de flexibilité** qui lui permet d'**ajuster** sa **complexité** selon le **contexte**, proposant un fonctionnement plus souple qu'un modèle entièrement figé.

1.5.1.3. Réduction de la Spécialisation Rigide

Dans un **réseau profond** standard comme un **CNN** dédié à la reconnaissance visuelle, la structure se présente souvent de manière **verticalisée**, où les premières couches extraient des **bords**, puis des formes plus élaborées, jusqu'à la classification finale. Cette disposition rend la **réutilisation** des *features* complexes si l'on souhaite aborder d'autres tâches ou modalités, à moins d'effectuer un **fine-tuning** parfois lourd. Les couches initiales sont fixées dans une fonction très spécifique, et la **coopération** entre différents blocs (par exemple la fusion de vision et d'audio) se voit

généralement cantonnée à un **niveau** imposé dans l’architecture ou à un **module** explicitement conçu pour la multimodalité.

Dans le cadre du **Deep Synergy Learning (DSL)**, la vision se veut plus **flexible**. Les **entités** composant le **Synergistic Connection Network (SCN)** incarnent chacune une **source d’information** ou un **bloc fonctionnel** susceptible d’être réutilisé dans d’autres contextes. La **synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$ détectée entre deux **entités** \mathcal{E}_i et \mathcal{E}_j ne se limite pas à une relation hiérarchique entre un étage “plus haut” ou “plus bas”. Toute entité peut, au fil des itérations, **établir un lien** avec une autre, dès lors que leur coopération génère un **gain** quantifié par la fonction S .

Un **exemple** concret peut être imaginé lorsqu’un **bloc** $\mathcal{E}_{\text{bords}}$, spécialisé dans la détection de **bords** visuels, observe soudain qu’il existe un **gain** d’associer ces **features** à un **bloc** $\mathcal{E}_{\text{audio}}$ focalisé sur les **fréquences** sonores, notamment quand des corrélations entre mouvements de lèvres et signaux vocaux se manifestent. Au sein d’un **CNN** traditionnel, l’**intégration** de telles informations se produirait typiquement à un **niveau tardif** ou via un **module spécialisé**, ce qui limite la portée de la coopération. Dans le **DSL**, au contraire, la liaison $\omega_{\text{bords}, \text{audio}}$ peut croître de manière **spontanée** si la **synergie** calculée (par exemple $\exp(-\|\mathbf{x}_{\text{bords}} - \mathbf{x}_{\text{audio}}\|^2/\sigma^2)$) s’avère **élevée** et dépasse un certain **seuil** ω_{\min} (section 1.4.5). Les deux blocs initient alors un **micro-cluster**, échangent leurs informations, et ainsi la **fusion** s’opère de manière autonome. Cette capacité à **tisser** des liens coopératifs entre entités diverge de la spécialisation rigide instaurée par des **pipelines** de couches fixes, et elle illustre la **souplesse** du DSL pour exploiter des **synergies** inattendues.

1.5.1.4. Facilité d’Intégration Multi-Entités

Dans la plupart des **réseaux neuronaux profonds** classiques, il est usuel de concevoir un **design** architectural bien particulier pour **fusionner** plusieurs entrées issues de différentes modalités, comme une **image**, un **texte** et un **signal audio**. On trouve ainsi des approches où un **CNN** traite l’image, un **RNN** (ou un **Transformers**) interprète le texte, puis une **couche de fusion** se situe au sommet pour combiner les vecteurs latents. Cette disposition requiert, dans son principe, une intervention humaine afin de spécifier à quel niveau et selon quelles règles la **fusion** se produit.

Dans le **Deep Synergy Learning (DSL)**, la perspective est nettement plus **flexible**. Chaque **modalité** s’incarne dans une ou plusieurs **entités** \mathcal{E}_{mod} dédiées, que l’on peut définir comme un bloc perceptif ou un ensemble de **features** cohérent. Les **liens** $\omega_{i,j}$ se **construisent** de manière autonome, selon la **synergie** mesurée entre les entités \mathcal{E}_i et \mathcal{E}_j . On peut exprimer cette dynamique à l’aide de la règle d’adaptation :

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)],$$

où η désigne le **taux d’apprentissage**, τ un **terme de régulation**, et $S(\mathcal{E}_i, \mathcal{E}_j)$ la **synergie** (section 1.4.4). Cette règle fait en sorte que toute **coopération** mutuellement avantageuse s’**intensifie** si elle contribue à un gain, alors que les associations moins pertinentes se **dissipent**. Dans un **contexte** multimodal, l’image, le texte et l’audio sont autant d’entités $\{\mathcal{E}_{\text{image}}, \mathcal{E}_{\text{texte}}, \mathcal{E}_{\text{audio}}\}$ susceptibles de découvrir des **corrélations** bilatérales ou trilatérales, sans qu’une **couche de fusion** imposée par un concepteur soit nécessaire.

Lorsque la **synergie** entre ces modalités se révèle conséquente, les pondérations $\omega_{\text{image, texte}}$, $\omega_{\text{image, audio}}$ ou $\omega_{\text{texte, audio}}$ croissent, formant de manière **spontanée** un **cluster** multimodal où les entités coopèrent directement. Cette **auto-organisation** se renforce si l'association apporte un bénéfice mesurable (par exemple, une performance améliorée sur une tâche de classification ou une plus grande robustesse). Inversement, si une modalité \mathcal{E}_{mod} n'apporte pas d'information utile ou se révèle trop bruitée, la **synergie** reste faible et la liaison ne s'établit pas, ou se rompt rapidement.

Une autre conséquence réside dans la **capacité** du réseau à **adapter** en continu les entités multimodales, autorisant la **formation** et la **dissolution** d'un **cluster** au gré du contexte. Cette propriété est en net contraste avec l'idée traditionnelle d'une couche de fusion figée où toutes les modalités se rejoignent selon un schéma unique. Le **DSL** propose donc une **répartition** dynamique de l'intégration multimodale, dans laquelle les **entités** se reconfigurent en fonction de l'**opportunité** d'apprentissage.

Au final, la **facilité d'intégration multi-entités** découle directement de la manière dont les **liaisons synergiques** sont définies et mises à jour. Chaque modalité, qu'elle concerne le **visuel**, le **textuel**, le **sonore** ou un **autre flux**, est **libre** de **découvrir** et de **former** des **connexions** directes avec n'importe quel **bloc** complémentaire, dès lors qu'une **valeur ajoutée** en résulte. Le **réseau** opère ainsi sous une forme d'**auto-organisation** (sections 1.4.3 et 1.4.5), permettant d'exploiter la diversité des **modalités** sans nécessiter le design d'un unique **module** de fusion, et garantissant une **adaptation** continue à la **variabilité** de l'environnement.

1.5.1.5. Apprentissage Continu

Dans de nombreux **réseaux neuronaux profonds**, l'entraînement se déroule dans un cadre essentiellement statique. On se contente généralement d'effectuer une rétropropagation sur un **ensemble de données** figé, et il s'avère difficile de réaliser un **apprentissage continu** sans retomber dans le phénomène de **catastrophic forgetting** au moment d'acquérir de nouvelles tâches ou de faire face à un changement de distribution. De plus, la structure même du réseau (couches, neurones) demeure immuable pendant tout l'apprentissage, ce qui limite considérablement la possibilité d'**évoluer** au fil du temps.

À l'inverse, le **Deep Synergy Learning (DSL)** s'appuie sur un **mécanisme** de **pondérations adaptatives** (discuté en section 1.4.5) dans lequel chaque liaison $\omega_{i,j}(t)$ peut se créer, se renforcer ou disparaître selon la **synergie** qu'elle procure. Cette propriété ouvre la voie à un **apprentissage continu** plus fluide, car le **réseau** peut accueillir de nouvelles **entités** $\mathcal{E}_{\text{nouv}}$ chaque fois qu'un nouveau flux de données ou une nouvelle modalité fait son apparition. Les **liaisons** $\omega_{\text{nouv},j}$ se forment si elles sont jugées bénéfiques, tandis que les anciens **clusters** jugés toujours utiles demeurent stables. Ainsi, l'**oubli** des acquis se trouve **atténué**, puisque les **connections** consolidées par le passé ne sont pas supprimées tant qu'elles conservent une synergie positive. Les entités correspondant aux tâches ou domaines précédents (sections 1.4.3 et 1.4.4) ne disparaissent pas, et les sous-groupes de coopération qui s'étaient formés peuvent persister ou se réactiver ultérieurement.

Une telle **flexibilité** convient particulièrement aux **environnements dynamiques**, où le **réseau** doit sans cesse **apprendre** de nouvelles classes d'objets, de nouvelles langues, ou s'adapter à des évolutions de contexte sans recommencer l'**entraînement** depuis zéro. Le **DSL** se distingue alors de la rigidité caractéristique d'un **réseau profond** standard, qui impose généralement de réinitialiser ou de "fine-tuner" un ensemble de couches fixes lors de l'introduction d'un nouveau domaine. La possibilité de **reconfigurer** la topologie et d'**absorber** ou de **cesser** d'utiliser certaines sources de données fait du **SCN** un cadre mieux adapté à un **apprentissage Lifelong** ou **continu**, tout en préservant la robustesse et la puissance d'un paradigme sub-symbolique.

Dans les sous-sections suivantes (1.5.2 à 1.5.7), nous approfondirons d'autres avantages majeurs, comme la **gestion naturelle de la multi-modalité** (1.5.2), la **flexibilité** face aux données partielles (1.5.3), l'**auto-évolution** (1.5.4), la **réduction** de la supervision (1.5.5), la **création** de représentations plus **riches** (1.5.6) et l'**intégration** de dimensions **symboliques** (1.5.7). L'ensemble consolidera l'idée qu'une **approche synergique** offre un panel d'atouts pour concevoir des systèmes apprenants plus **évolutifs**, **modulaires** et **généraux** que ne le permettent les architectures entièrement figées du deep learning traditionnel.

1.5.2. *Gestion Naturelle de la Multi-modalité*

Une des motivations fortes qui a conduit à l'élaboration du **Deep Synergy Learning (DSL)** est la possibilité d'**intégrer** plusieurs types de données (*images, sons, textes, capteurs variés, etc.*) de manière qu'elles **coopèrent** au lieu de simplement être fusionnées dans une couche dédiée. Dans les réseaux neuronaux profonds classiques, la multi-modalité impose souvent de **concevoir** des modules spécifiques (p. ex. un CNN pour l'image, un RNN ou un Transformer pour le texte, un autre réseau pour l'audio), puis de **fusionner** ces modules dans une partie supérieure de l'architecture. Cette approche, bien qu'efficace dans de nombreux cas, demeure relativement **rigide**. La fusion s'effectue à un niveau **préétabli** et évolue peu en fonction du **contexte** ou de la **dynamique interne** du réseau.

Le **DSL**, à l'inverse, favorise une **intégration plus organique**. Chaque modalité est représentée par une ou plusieurs **entités** \mathcal{E}_{mod} , qui ajustent **spontanément** leurs connexions en fonction de la **synergie** détectée. Elles peuvent ainsi **établir** ou **rompre** des liens avec des entités d'autres modalités, garantissant une flexibilité et une adaptation continue du réseau. Cette sous-section (1.5.2) détaille :

- **Pourquoi** la multi-modalité est "naturellement" absorbée par le DSL,
- Comment les entités **visuelles**, **auditives**, **textuelles** (par ex.) peuvent **former** des clusters multimodaux,
- Quels **avantages** cette gestion apporte (robustesse, découvertes de liens inattendus, etc.).

1.5.2.1. Les Entités comme Vecteurs ou Blocs Multimodaux

Dans le **Deep Synergy Learning (DSL)**, chaque **entité** \mathcal{E}_i peut se spécialiser dans un type de données particulier ou combiner déjà plusieurs **descripteurs**. Il n'y a pas l'obligation, comme dans

un réseau multimodal classique, d'assembler deux flux (image, audio, etc.) au niveau d'une couche imposée ou dans des branches séparées. Au lieu de cela, toutes les entités coexistent dans un **même "espace"** ou **graphe** $\{\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_n\}$. On y retrouve des entités $\mathcal{E}_{\text{image}}$, $\mathcal{E}_{\text{audio}}$, $\mathcal{E}_{\text{texte}}$, ou des blocs déjà **multimodaux** combinant différents **features**. Il n'existe pas de **pyramide** de couches fixes, comme un **CNN** qui détecterait d'abord les bords, puis les textures, puis les objets. Au contraire, le **DSL** insiste sur le principe d'**auto-organisation**, faisant que toute entité \mathcal{E}_i peut se **lier** à toute autre \mathcal{E}_j dès lors qu'une **synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$ significative est détectée. La mise à jour de la pondération $\omega_{i,j}(t)$ suit la règle discutée en section 1.4.5, et si la coopération apporte un gain, on obtient un renforcement de la liaison $\omega_{i,j}$. Ainsi, une entité de type *texte* \mathcal{E}_{txt} n'est pas contrainte de se fusionner uniquement avec une entité *image* \mathcal{E}_{img} à un étage supérieur prédéfini. La **fusion** peut se produire **à tout moment**, entre n'importe quelles entités, dès lors qu'une **valeur ajoutée** est détectée. Cette démarche favorise la **formation** naturelle de **clusters** multimodaux là où c'est nécessaire, et permet aux entités déjà mixtes (par exemple combinant audio et vision) de se relier également à d'autres (texte, capteurs...), rendant le réseau plus **souple** et mieux à même de gérer la **pluralité** des sources de données.

1.5.2.2. Formation de Clusters Multimodaux

Le **Deep Synergy Learning (DSL)** emploie une règle de mise à jour décrite dans la section 1.4.5, qui se formalise ainsi :

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)].$$

Lorsque, par exemple, une **entité visuelle** $\mathcal{E}_{\text{visuelle}}^a$ (extraction d'images) et une **entité auditive** $\mathcal{E}_{\text{auditive}}^b$ (analyse du spectre sonore) constatent une **synergie** satisfaisante, la liaison $\omega_{a,b}$ se **renforce** progressivement. Les sections 1.4.3 et 1.4.4 évoquent déjà ce principe de consolidation, où toute **coopération** profitable se traduit par une augmentation de la pondération. Au fil des itérations, d'autres entités, qu'elles soient textuelles ou associées à une autre modalité, peuvent se **greffer** à ce duo si elles y perçoivent-elles aussi un **gain**. S'organise alors un **cluster** multimodal $\{\mathcal{E}_{\text{vis}}, \mathcal{E}_{\text{aud}}, \mathcal{E}_{\text{txt}}\}$, destiné, par exemple, à la reconnaissance d'événements conjoints dans une vidéo, en y associant le son et des sous-titres.

Un exemple important survient lors de l'analyse synchronisée **vidéo + audio**. Une entité \mathcal{E}_{CNN} , spécialisée dans la détection d'objets ou de mouvements visuels, et une entité $\mathcal{E}_{\text{Audio}}$, orientée vers les fréquences sonores, découvrent une **corrélation** entre le contenu visuel (mouvements de l'orateur) et des sons spécifiques. Si la **fonction de synergie** $S(\mathcal{E}_{\text{CNN}}, \mathcal{E}_{\text{Audio}})$ affiche une valeur élevée, la pondération $\omega_{\text{CNN}, \text{Audio}}$ croît. Il en résulte la formation d'un **sous-réseau** constitué de ces entités, apte à identifier les vidéos où un son précis se déclenche simultanément avec un geste ou un mouvement labial.

Un autre scénario se rencontre dans la **fusion texte + image**. Si une entité $\mathcal{E}_{\text{visuel}}$ extrait les objets visibles dans une scène et qu'une entité $\mathcal{E}_{\text{langage}}$ génère des phrases descriptives (via un modèle de type RNN ou Transformers), la **co-information** qu'elles partagent peut montrer qu'en les associant, on obtient une légende automatique bien plus fidèle. Cette plus-value se reflète dans la

liaison $\omega_{\text{visuel}, \text{langage}}$, qui se **solidifie** de manière autonome. Le **cluster** $\{\mathcal{E}_{\text{visuel}}, \mathcal{E}_{\text{langage}}\}$ se spécialise alors dans la génération de légendes, sans imposer de couche de fusion prédéterminée.

On peut enfin imaginer un **cluster hybride** regroupant à la fois la “capture de mouvement” via des capteurs inertiels, la “reconnaissance d’images” au moyen d’une entité visuelle, et l’“analyse du spectre sonore”. Ce **triplet** se cristallise dès lors qu’il existe une forte complémentarité pour identifier des événements qui combinent aspects visuels, audio, et déplacements corporels. Le **SCN** encourage ainsi la création de tels **clusters** dès que les synergies apparaissent bénéfiques, ce qui permet au réseau de s’adapter aux environnements multimodaux et de tirer le meilleur parti de chacune des sources disponibles.

1.5.2.3. Avantages de la Co-Organisation Multimodale

Dans un **réseau neuronal** habituel, la fusion des différents **flux** (par exemple, l’image et le son) est généralement décidée à l’avance, souvent dans les étages supérieurs ou dans une **couche intermédiaire** prédéfinie. Le **Deep Synergy Learning (DSL)** adopte une approche différente. Chaque **entité** \mathcal{E}_i peut **rechercher** spontanément des partenaires synergiques, **évaluer** si la coopération améliore les performances ou génère un gain en information mutuelle, et **établir** des liaisons directes $\omega_{i,j}$ sans qu’une couche de fusion prédéfinie ne soit requise. Cette démarche favorise l’apparition de **combinaisons** inhabituelles mais profitables, comme l’association de l’analyse audio, de la mesure de température et de l’intensité lumineuse, si un cluster pertinent émerge.

La **règle** de mise à jour

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)]$$

permet à chaque entité de constater, de façon autonome, si sa **synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$ s’avère suffisamment élevée pour justifier une **collaboration**. Ainsi, il n’existe pas d’étage unique où l’on agrégerait tous les flux, mais plutôt un **réseau** dont la topologie évolue pour intégrer toute coopération bénéfique.

Lorsqu’une modalité est **perturbée** ou moins fiable (par exemple l’audio, soumis à un **bruit** important ou une panne de capteur), la synergie associée $\omega_{\text{aud}, .}$ s’affaiblit, car les entités n’y trouvent plus de **valeur ajoutée**. Le **cluster** multimodal se réorganise donc en renforçant d’autres connexions, comme celles reliant la vision, le texte ou des capteurs alternatifs. Le système gagne ainsi en **résilience**. Il ne repose pas sur un pipeline figé ni sur un module dédié à chaque modalité, mais s’**autorégule** en continu pour maintenir une **robustesse** globale, s’adaptant aux changements de contexte et compensant d’éventuelles défaillances de certaines sources.

1.5.2.4. Éléments Mathématiques : Synergie “Multi-Modal”

Dans le **Deep Synergy Learning (DSL)**, la mesure de **synergie** entre différentes **modalités** (comme l’audio, la vision ou le texte) peut prendre plusieurs formes selon le type de données ou la définition souhaitée de la **co-opération**. Deux grandes approches sont envisagées. La première

repose sur l'utilisation de la **co-information** (ou multi-information) issue de la théorie de l'information, tandis que la seconde exploite une **distance** ou une **similarité** inter-modale définie dans un **espace latent commun**.

A. Mesure de co-information

Lorsque l'on manipule un ensemble de modalités $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_m$, il est possible d'estimer une **co-information** ou **mutual information** globale :

$$S(\mathbf{X}_1, \dots, \mathbf{X}_m) = \text{MI}(\mathbf{X}_1, \dots, \mathbf{X}_m).$$

Dans certains cas, la **Partial Information Decomposition (PID)** permet de séparer la part strictement *synergique* de la part *redondante* ou *unique*. Dans le **DSL**, on s'intéresse à l'**analyse n-aire** au sens où si la **conjonction** de deux (ou plus) modalités, par exemple $\mathbf{X}_{\text{visuelle}}$ et $\mathbf{X}_{\text{auditive}}$, produit une **information** qu'aucune des modalités ne détenait isolément, alors on en conclut un **gain** pour la **synergie**. Ce **gain** se traduit par un **renforcement** des pondérations. Lorsque l'association d'un flux visuel et d'un flux audio se révèle bénéfique, la liaison $\omega_{\text{vis}, \text{aud}}$ se renforce conformément à la règle de mise à jour décrite en **section 1.4.5**. L'**auto-organisation** du réseau valorise ainsi la coopération entre modalités lorsqu'elle apporte un surcroît d'information ou de performance.

B. Distance ou Similarité inter-modale

En alternative à la co-information, on peut quantifier la **synergie** via une **distance** (ou **similarité**) entre les modalités, à condition de se situer dans un **espace latent** commun. Dans la pratique, cela implique généralement les étapes suivantes :

1. Embeddings partagés. On forme une représentation vectorielle \mathbf{z}_{vis} pour la partie **visuelle** et \mathbf{z}_{aud} pour la partie **audio**, toutes deux dans \mathbb{R}^d . Ces embeddings peuvent provenir d'un réseau d'apprentissage dédié ou être appris conjointement.

2. Distance ou similarité. On calcule une **norme** $\|\mathbf{z}_{\text{vis}} - \mathbf{z}_{\text{aud}}\|$ ou une **similarité cosinus** $\langle \mathbf{z}_{\text{vis}}, \mathbf{z}_{\text{aud}} \rangle$. Cette quantité rend compte de la **proximité** ou de la **complémentarité** entre les deux modalités une fois projetées dans l'espace latent.

3. Définition de la synergie. En inversant ou en normalisant la distance, on obtient un **score de coopération** :

$$S(\mathbf{z}_{\text{vis}}, \mathbf{z}_{\text{aud}}) = \exp(-\|\mathbf{z}_{\text{vis}} - \mathbf{z}_{\text{aud}}\|^2 / 2\sigma^2) \quad \text{ou} \quad 1 / (1 + \|\mathbf{z}_{\text{vis}} - \mathbf{z}_{\text{aud}}\|^2),$$

ou encore en recourant à une **similarité** normalisée dans $[0,1]$. Cette mesure, insérée dans la formule de mise à jour des pondérations $\omega_{i,j}(t+1) = \omega_{i,j}(t) + \dots$, renseigne sur la **pertinence** qu'il y a à **connecter** et **coordonner** ces deux modalités.

Synthèse et implication dans le DSL

Le **DSL** n'impose pas de définition unique de la synergie ; toute **fonction** mesurant la **coopération** peut s'appliquer. La **co-information** met en évidence des **complémentarités** complexes (non linéaires), en évaluant l'information que seules deux (ou plusieurs) modalités, **combinées**, parviennent à extraire. De son côté, la **distance** (ou la **similarité**) dans un **espace latent** offre un

calcul plus direct et souvent moins coûteux, permettant d'établir rapidement une **coopération** si les embeddings s'avèrent proches ou orientés de façon compatible.

Dans tous les cas, le **DSL** conserve son principe fondamental. La **synergie inter-modale** joue le rôle d'un **signal**, suggérant que certaines entités multimodales bénéficieraient d'une **connexion renforcée**, favorisant ainsi la formation d'un **cluster dédié**. Cette logique s'applique à deux entités (binaire), mais peut s'étendre à des groupes plus vastes (section 1.4.7) lorsque plusieurs **modalités** s'enrichissent mutuellement. Dans une perspective **multimodale**, on voit ainsi se constituer des **clusters** audio-vidéo-texte, ou audio-vision-capteurs, sans qu'un concepteur doive imposer explicitement à quel niveau se fait la **fusion**. Le **SCN** se reconfigure et s'**auto-organise**, renforçant les **liens** rentables au sens de la synergie globale.

1.5.2.5. Perspectives : Plus de Fluidité, Plus de Découvertes

Dans un **Deep Synergy Learning (DSL)**, la fusion multimodale n'est pas confinée à une couche ou un module imposé à l'avance. Elle repose sur l'aptitude de chaque **entité** à forger des liens **synergiques** dès lors que la coopération procure un **gain**. Cette configuration autorise plusieurs évolutions particulièrement souples.

Un premier aspect concerne la **création spontanée** de **clusters multimodaux**. Lorsqu'un événement sollicite simultanément divers flux, par exemple un son particulier, un motif visuel distinctif et un mot-clé textuel, le **DSL** favorise la formation d'un **cluster** auto-organisé combinant toutes ces sources. Il n'est pas nécessaire de prédéfinir, lors de la conception du réseau, à quel niveau ou dans quel sous-ensemble la fusion doit se produire. Le **Synergistic Connection Network** s'**auto-organise** en **détectant** les entités pertinentes et en **renforçant** leurs liaisons $\omega_{i,j}$ lorsque la **synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$ est élevée. Cette auto-organisation, évoquée en section 1.4.3, rend la fusion bien plus fluide et adaptative.

Un deuxième enjeu apparaît lorsque des **modalités** ou sources de données nouvelles se présentent. Si un capteur inédit vient produire un nouveau flux, il suffit d'introduire une entité $\mathcal{E}_{\text{capteur}}$ correspondante. Les pondérations $\omega_{\text{capteur},x}$ se mettent alors à jour selon la même loi que précédemment (section 1.4.5). Si les interactions avec des entités déjà existantes procurent un réel **bénéfice** à l'apprentissage ou à la représentation, la **synergie** grimpe ; sinon, le lien demeure faible et finit par s'éteindre. L'essentiel réside dans l'absence de nécessité de **reconstruire** l'architecture globale. Le réseau s'**auto-adapte**, formant et supprimant des connexions en fonction des **opportunités** offertes par le nouveau flux.

Un troisième point souligne la **réutilisation** des entités spécialisées. Dans un paradigme classique, un module dédié, par exemple à l'analyse de fréquences sonores, risque de ne pouvoir servir qu'à une tâche unique (identification de parole ou détection de musique). Au contraire, dans le **DSL**, une entité $\mathcal{E}_{\text{audio}}$ capable de détecter des éléments discriminants dans le spectre sonore peut être **sollicitée** par plusieurs **clusters** simultanément. Elle peut ainsi intégrer un sous-réseau dédié à la reconnaissance de mots clés, un autre spécialisé dans la détection de bruits inhabituels, ou encore un qui associe des fragments de parole aux mouvements des lèvres. Chaque fois que la **synergie** le justifie, la liaison $\omega_{\text{audio},...}$ se renforce dans le cluster concerné, sans imposer de duplication de la fonctionnalité ni de modification d'une couche fixe.

L'ensemble de ces mécanismes assure une **fluidité** et un **potentiel d'exploration** supérieurs par rapport aux modèles multimodaux figés. Le **DSL** favorise l'**émergence** et la **dissolution** dynamiques de clusters multimodaux, l'**intégration** graduelle de modalités nouvelles, et la **mutualisation** de blocs spécialisés au service de plusieurs tâches. Cette architecture, loin d'être imposée, se déploie au fil des itérations par simple adaptation des pondérations $\{\omega_{i,j}(t)\}$, conduisant à une plus grande **richesse** de découvertes et de combinaisons entre entités.

1.5.3. *Flexibilité vis-à-vis des Données Incomplètes ou Bruitées*

Outre la capacité à gérer de multiples flux (voir 1.5.2), l'un des atouts majeurs du **Deep Synergy Learning (DSL)** réside dans sa **tolérance** accrue face aux **données incomplètes** ou **fortement bruitées**. Dans les approches d'apprentissage traditionnelles (réseaux neuronaux profonds inclus), on cherche souvent à **normaliser** ou **compléter** les données manquantes, voire on opte pour un **prétraitement** lourd afin de filtrer les bruits. Le DSL, par sa **structure auto-organisée**, autorise un réseau à composer **localement** avec les lacunes et à **réajuster** ses connexions de manière à s'appuyer sur les sources les plus fiables.

Cette section (1.5.3) met en avant les principes qui confèrent au DSL sa **robustesse** et sa **souplesse** quand on fait face à des entrées partielles, hétérogènes ou bruitées.

1.5.3.1. *Rôle de l'Auto-Organisation dans la Gestion du Bruit*

Dans le **Deep Synergy Learning (DSL)**, les entités $\{\mathcal{E}_1, \dots, \mathcal{E}_n\}$ s'**agglomèrent** en **clusters** (tels que décrits en section 1.4.3) lorsque leurs synergies bilatérales $S(\mathcal{E}_i, \mathcal{E}_j)$ atteignent un certain **niveau**. L'auto-organisation agit alors comme un **mécanisme** filtrant. Si une entité \mathcal{E}_k se révèle trop **bruyante**, ou plus généralement peu profitable (synergie insuffisante, voire négative, avec les autres), ses connexions ω_k , **dépérissent** naturellement, selon la **mise à jour** :

$$\omega_{k,j}(t+1) = \omega_{k,j}(t) + \eta [S(\mathcal{E}_k, \mathcal{E}_j) - \tau \omega_{k,j}(t)].$$

Si la synergie $S(\mathcal{E}_k, \mathcal{E}_j)$ reste **faible** ou négative, la liaison $\omega_{k,j}$ diminue itération après itération jusqu'à se rapprocher d'un **seuil** ω_{\min} en deçà duquel elle est jugée **inexistante**. L'entité \mathcal{E}_k tend dès lors à **s'isoler** de la structure principale et ne contribue plus de manière significative aux décisions ou au regroupement de clusters, limitant l'**impact** du bruit ou des mesures peu fiables sur l'ensemble du **Synergistic Connection Network**. Cette propriété confère au **DSL** une **robustesse** importante, car il n'est pas nécessaire d'éliminer d'emblée les données incertaines. C'est la **dynamique** des pondérations qui réalise ce filtrage de manière **locale** et **autonome**, en ajustant les connexions selon la synergie détectée.

De façon analogue, une entité ne possédant que des **données incomplètes** (par exemple, un capteur qui ne fonctionne que par intermittence) peut tout de même établir des **liens forts** avec d'autres entités si la partie (même partielle) de ses observations est **pertinente** et apporte un gain synergique réel. L'apprentissage ne rejette pas d'emblée les capteurs partiellement défectueux, mais évalue la **valeur ajoutée** qu'ils apportent, comme expliqué dans la section 1.4.4 sur la définition de S . En fonction de cette évaluation, il **renforce** ou **dissout** les liaisons selon leur pertinence. Le **DSL** teste

ainsi localement l'**efficacité** des données partielles et, grâce à la dynamique adaptative de $\omega_{k,j}(t)$, il détermine si une entité doit être maintenue dans un cluster ou rester en marge du réseau.

Ce procédé évite la nécessité d'une **stratégie** spécifique pour rejeter le **bruit** ou manipuler les **données manquantes**. Le **système** s'auto-régularise, en quelque sorte, grâce à la diminution des pondérations non rentables, ce qui se traduit par une **atténuation** de l'influence des sources imprécises. L'**auto-organisation** réalise donc une forme de **gestion du bruit** distribuée, où chaque entité (ou module) se maintient ou s'éclipse selon la **synergie** réellement engendrée avec les autres composantes du **SCN**, évitant ainsi de nuire à la **performance** globale.

1.5.3.2. Interprétation dans un Cadre Bruité

Lorsqu'une entité \mathcal{E}_k se trouve exposée à un **bruit** important dans ses représentations \mathbf{x}_k , sa **similarité** ou son **information mutuelle** avec les autres entités demeure faible ou fluctuante. On peut considérer par exemple la fonction

$$S(\mathcal{E}_k, \mathcal{E}_j) = \frac{1}{1 + \|\mathbf{x}_k - \mathbf{x}_j\|^2}.$$

Si \mathbf{x}_k varie de manière aléatoire (bruit élevé), alors $\|\mathbf{x}_k - \mathbf{x}_j\|$ reste souvent grande ou instable, réduisant ainsi la valeur de $S(\mathcal{E}_k, \mathcal{E}_j)$. En se référant à la loi de mise à jour des pondérations (section 1.4.5), on voit que

$$\omega_{k,j}(t+1) = \omega_{k,j}(t) + \eta [S(\mathcal{E}_k, \mathcal{E}_j) - \tau \omega_{k,j}(t)].$$

La valeur de $\omega_{k,j}(t)$ décroît donc si la **synergie** $S(\mathcal{E}_k, \mathcal{E}_j)$ ne compense pas suffisamment le terme de régulation $\tau \omega_{k,j}(t)$. Cette **diminution** progressive des liaisons $\omega_{k,j}$ conduit l'entité \mathcal{E}_k à s'**isoler** si le bruit la rend incohérente vis-à-vis des entités du cluster. Il n'est pas nécessaire d'ajouter un algorithme spécifique pour exclure certaines entités, car le **DSL** prend en charge ce processus de manière **auto-organisée**, en laissant le lien s'affaiblir jusqu'à potentiellement s'éteindre lorsque

$$\omega_{k,j}(t) < \omega_{\min}.$$

Dans l'hypothèse où, plus tard, \mathcal{E}_k retrouve une plus grande **fiabilité** (par exemple si un paramètre interne $\mathbf{s}_k(t)$ se recalibre, ou si le flux de données s'assainit), on observe que la **distance** $\|\mathbf{x}_k - \mathbf{x}_j\|$ peut redevenir modérée, augmentant $S(\mathcal{E}_k, \mathcal{E}_j)$. La mise à jour $\omega_{k,j}(t+1)$ peut alors reprendre une tendance à la hausse, recréant ou solidifiant des liaisons $\omega_{k,j}$. Cette **réversibilité** illustre l'une des forces du **Synergistic Connection Network**. Une entité qui s'était tenue à l'écart en raison d'un bruit excessif peut retrouver **sa place** dès que ses **observations** redeviennent pertinentes, sans qu'une restructuration globale du réseau soit nécessaire. Elle réintègre alors le **cluster**, démontrant ainsi la **capacité** du **DSL** à s'**adapter** en continu aux fluctuations de la qualité des données.

1.5.3.3. Données Partielles : Complétion Progressive via Synergie

Dans de nombreux **scénarios** pratiques, certaines entités ne disposent que d'une **fraction** de leurs attributs. Par exemple, dans un cadre médical, un **patient** peut présenter des **données** incomplètes en raison de tests sanguins manquants, d'une imagerie non réalisée ou d'un historique médical lacunaire. Les approches neuronales classiques exigent souvent une **imputation** ou un **remplissage** préalable des valeurs absentes, ce qui nécessite des hypothèses ou des méthodes de substitution. Il n'est pas rare non plus que l'on décide simplement d'**écarter** les exemples incomplets pour ne pas perturber l'apprentissage.

Dans le **Deep Synergy Learning (DSL)**, une telle entité — par exemple $\mathcal{E}_{\text{patient}}$ — est autorisée à **coopérer** partiellement avec d'autres entités. Si l'entité patient ne possède que certaines **mesures** (tests sanguins, variables cliniques) mais en ignore d'autres, elle peut néanmoins évaluer sa **synergie** avec, par exemple, d'autres patients $\mathcal{E}'_{\text{patient}}$ ou des **variables globales** (moyennes de cohortes, règles médicales) en se basant sur les **dimensions** ou **attributs** qu'elle détient réellement. Une fonction de synergie $S(\mathcal{E}_k, \mathcal{E}_j)$ adaptée aux **dimensions communes** ou à une distance partielle peut alors guider la mise à jour de la pondération :

$$\omega_{k,j}(t+1) = \omega_{k,j}(t) + \eta [S(\mathcal{E}_k, \mathcal{E}_j) - \tau \omega_{k,j}(t)].$$

Si, malgré des données incomplètes, l'entité \mathcal{E}_k dégage une **valeur ajoutée** non négligeable pour le réseau (par exemple, la portion de tests sanguins disponibles est très informative), ses **liaisons** $\omega_{k,j}$ avec d'autres entités \mathcal{E}_j se **consolident** progressivement. Cela peut se traduire par la **constitution** d'un **cluster** rassemblant d'autres patients, ou bien des entités spécialisées (un bloc fonctionnel détectant une tendance clinique). L'**incomplétude** des données n'empêche pas la **coopération**, car seul compte le **gain** mesuré sur les attributs effectivement partagés.

Au fur et à mesure que l'entité \mathcal{E}_k acquiert (ou reconstitue) de nouveaux attributs manquants, il est envisageable que sa **distance** ou sa **similarité** avec d'autres entités évolue, rehaussant la pondération $\omega_{k,j}$. Le réseau en tire un bénéfice, car la réévaluation de la synergie intègre cette **nouvelle information**, facilitant ainsi le **réajustement** des liens. Cela permet d'éviter l'**exclusion** systématique des cas incomplets ou le recours à une **imputation** globale. Au lieu de cela, le **DSL** se contente de **tester** la **collaboration** sur la partie des données effectivement disponible. Cette logique de **liaison partielle** confère au **DSL** une **souplesse** précieuse pour manipuler des **jeux de données** irréguliers, tout en conservant la capacité d'un **regroupement** (cluster) auto-organisé autour des **similarités** ou **complémentarités** réellement détectées.

1.5.3.4. Exemples Concrets et Bénéfices

Dans un **environnement industriel**, certains capteurs (de température, de pression...) se révèlent souvent **défaillants** ou trop **bruités**. Dans un **Deep Synergy Learning (DSL)**, les entités correspondantes $\mathcal{E}_{\text{temp}}$ ou $\mathcal{E}_{\text{press}}$ voient leurs **pondérations** $\omega_{\text{temp},j}$, $\omega_{\text{press},j}$ diminuer dès lors que la **synergie** avec d'autres capteurs (comme $\mathcal{E}_{\text{vibration}}$ ou $\mathcal{E}_{\text{débit}}$) ne justifie plus la coopération. Le **cluster** principal, dédié par exemple à la **détection d'anomalies**, se recentre naturellement sur les capteurs les plus fiables. Il n'est pas nécessaire d'exclure explicitement les capteurs défaillants, car ceux-ci s'**isolent d'eux-mêmes** lorsque leurs liaisons $\omega_{i,j}$ deviennent insignifiantes. Si, par la

suite, un capteur perturbé se **rétablit** (ou se recale), les pondérations se remettent à monter, lui permettant de **réintégrer** le cluster d'intérêt. Ce **mécanisme** procure une **robustesse** naturelle face aux dérives ponctuelles de certains capteurs.

Dans un autre contexte, celui d'une **base utilisateur** (profilage, recommandation), il est courant que les **profils** ne renseignent pas l'intégralité des informations attendues. En pratique, un **pipeline** neuronal classique suppose souvent une **imputation** des champs manquants ou l'exclusion des données incomplètes. Le **DSL**, au contraire, autorise l'insertion d'entités représentant « Utilisateur U1 » avec un vecteur \mathbf{x}_{U1} partiel, ainsi que d'entités « Contenu C1, C2 » définies par divers attributs. La **synergie** $S(\mathcal{E}_{U1}, \mathcal{E}_{C2})$ tient compte des attributs effectivement présents. Si, même partiellement, l'utilisateur U1 et un contenu C2 trouvent un **recoupement** significatif, la pondération $\omega_{U1, C2}$ s'accroît sans imposer que tous les champs soient renseignés. Le **réseau** s'**auto-structure** donc malgré l'incomplétude, sans avoir à multiplier les procédures d'imputation ou à écarter systématiquement les exemples incomplets. Cela permet de **préserver** une large portion d'information disponible et d'exploiter la **valeur ajoutée** dès qu'elle se présente.

1.5.3.5. Comment le DSL Surpasse la Rigidité des Réseaux Profonds

Les **réseaux neuronaux profonds** présentent une sensibilité notable aux **incohérences** ou aux cas situés hors de la distribution rencontrée lors de l'entraînement. Si, par exemple, un flux de capteurs n'a jamais inclus des observations particulièrement bruitées, le **modèle** peine à gérer ces perturbations lorsqu'elles apparaissent. En revanche, le **Deep Synergy Learning (DSL)** se caractérise par une **auto-régulation**. Lorsqu'une entité \mathcal{E}_k se montre trop incertaine (synergie très basse ou négative avec d'autres entités), les pondérations $\omega_{k,j}(t)$ s'amoindrissent, et la contribution de \mathcal{E}_k se trouve mécaniquement **limitée** ou mise à l'écart, au moins de manière temporaire. Le **SCN** continue malgré tout de fonctionner avec les entités restantes, potentiellement plus fiables.

Dans bien des cas classiques, on recourt à un **modèle d'imputation** (pour combler les données manquantes) ou à des heuristiques rigides visant à traiter le **bruit**. Dans le **DSL**, à l'inverse, chaque entité **évalue** localement la **synergie** possible, sans qu'un algorithme unique de “remplissage” s'impose. Cette démarche revient à laisser chaque **liaison** $\omega_{k,j}$ apprendre la **compatibilité** ou l'**incompatibilité** entre \mathcal{E}_k et \mathcal{E}_j , de sorte que, lorsque le bruit ou l'incohérence survient, la pondération se détériore progressivement, isolant l'entité problématique du **cluster** principal.

Si la **qualité** des données associées à une entité fluctue dans le temps — qu'il s'agisse d'un capteur dont la précision varie ou d'une source intermittente de bruit — la **dynamique** d'auto-organisation gère cette évolution. Les **liaisons** $\omega_{k,j}$ s'étirent ou se resserrent selon la **synergie** présente, et aucune **réinitialisation** globale ni refonte de l'**architecture** n'est requise. Le réseau d'entités conserve sa topologie, ajustant simplement les **connexions** nécessaires. Cette **flexibilité** évite l'extrême rigidité à laquelle aboutit souvent un réseau profond traditionnel, où l'adaptation à de nouveaux modes de bruit ou d'incomplétude implique souvent un **réentraînement** sur un large jeu de données.

1.5.4. Potentiel d'Auto-Évolution et d'Adaptation Continue

Un des aspects les plus novateurs du **Deep Synergy Learning (DSL)** concerne sa capacité à **évoluer** au fil du temps, sans nécessairement repasser par une phase d'entraînement globale et figée. Plutôt que de geler l'architecture après avoir ajusté quelques poids, le DSL propose une dynamique **en continu** où la structure interne (liens synergiques, clusters, etc.) se reconfigure régulièrement. Cette **auto-évolution** rend le réseau apte à faire face à des changements de distribution (domain shift), à l'apparition de nouvelles entités (nouvelles sources de données) ou à la nécessité d'exploiter un flux continu de données (*streaming data*). Dans cette section (1.5.4), nous examinons :

- Les principes d'**auto-évolution** du DSL,
- Comment cette dynamique repose sur l'**adaptation continue** des pondérations (déjà décrite en partie en 1.4.5),
- Les bénéfices en termes de **Lifelong Learning** (apprentissage tout au long de la vie) et de **plasticité** comparable à des systèmes vivants.

1.5.4.1. L'Auto-Organisation comme Moteur d'Évolution

Dans le **Deep Synergy Learning (DSL)**, la **pondération** $\omega_{i,j}(t)$ reliant deux entités \mathcal{E}_i et \mathcal{E}_j suit la règle de mise à jour décrite en section 1.4.5. La pondération se met à jour selon

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)],$$

où $S(\mathcal{E}_i, \mathcal{E}_j)$ correspond à la **synergie** (distance, similarité, co-information, etc.). Si cette **synergie** demeure suffisamment **positive**, la liaison $\omega_{i,j}$ se **renforce** ; dans le cas contraire, elle **décroît** et peut passer en deçà d'un **seuil** (voir ω_{\min} en section 1.4.5). Un lien sous ce seuil est généralement jugé **inexistant**, reflétant un **désintérêt** ou une **incohérence** mutuelle.

Ce **principe d'évolution dynamique** des connexions fait émerger des **clusters**. Deux entités qui constatent un **gain mutuel** se **rapprochent**, et des sous-groupes peuvent dès lors **fusionner** si la synergie inter-clusters croît au fil des itérations. À l'inverse, un **cluster** se **scinde** quand des sous-groupes décelant davantage de synergie interne choisissent de s'éloigner du groupe initial, leurs connexions inter-groupes ω s'abaissant peu à peu faute de bénéfice collectif.

Ce **mécanisme** confère au **DSL** une **dimension** "vivante" : le **Synergistic Connection Network** évolue en continu, s'**adapte** à l'arrivée de nouvelles données ou à la modification d'une source d'information, tout en **préservant** les clusters confirmés et en **permettant** l'émergence de combinaisons inédites. À aucun moment, un **réentraînement** exhaustif du réseau n'est nécessaire. La **dynamique locale** des pondérations $\{\omega_{i,j}\}$ permet une **réorganisation progressive**, en ajustant les **liaisons** selon l'opportunité détectée dans les données, que ce soit pour renforcer des connexions existantes ou en dissoudre d'autres.

1.5.4.2. Lifelong Learning et Intégration de Nouvelles Entités

Dans de nombreux environnements dynamiques (robotique, systèmes d'information évolutifs, etc.), il est courant de rencontrer des **données** ou des **entités** inédites au fil du temps. Les réseaux neuronaux profonds classiques rencontrent plusieurs difficultés dans ce contexte. L'ajout d'une nouvelle fonctionnalité peut nécessiter la création d'un **module annexe** et un **réentraînement**, parfois complet ou via un **fine-tuning**, ce qui comporte le risque de perturber l'**équilibre établi** du modèle. Il arrive aussi que l'on doive réorganiser l'architecture en profondeur, ce qui affecte la stabilité du modèle.

Le **Deep Synergy Learning (DSL)** propose une alternative plus organique. Pour incorporer une **entité** $\mathcal{E}_{\text{nouvelle}}$ correspondant à un nouveau flux de données, on se contente de l'**introduire** dans le graphe existant. Les pondérations la reliant à d'autres entités $\{\omega_{\text{nouvelle},i}\}$ s'ajustent alors en fonction de la synergie détectée. Si une **coopération** apparaît avantageuse, ces liens augmentent et conduisent à l'**insertion** naturelle de l'entité dans un cluster. Il n'est donc pas nécessaire de reconfigurer l'ensemble de l'architecture, car seules les **connexions pertinentes** se renforcent de manière autonome.

Ce mécanisme permet une **évolution progressive**. Les **clusters** déjà établis maintiennent leurs **liaisons** tant qu'ils apportent un **gain**, garantissant ainsi la stabilité de leurs synergies internes. Le réseau n'est pas obligé de sacrifier ses acquis pour introduire de nouvelles entités. Ainsi, on **atténue** le phénomène de "catastrophic forgetting" caractéristique des modèles neuronaux, lesquels risquent d'effacer de vieilles connaissances lorsqu'on leur enseigne de nouvelles tâches. Le **DSL**, au lieu de **remplacer** brutalement un jeu de poids par un autre, **réorganise** progressivement les connexions de manière **adaptative locale**. Il conserve la pertinence des **clusters existants** tout en intégrant de nouvelles entités au réseau lorsqu'elles apportent une **valeur ajoutée**.

1.5.4.3. Adaptation Continue aux Changements de Distribution

Dans un **réseau** neuronal traditionnel, lorsqu'apparaît un **changement de distribution** dans les données (par exemple l'évolution des conditions de capteurs ou l'apparition de nouvelles caractéristiques), il est souvent indispensable de **réentraîner** ou de **fine-tuner** le modèle, ce qui peut provoquer un ajustement délicat, risqué pour la stabilité ou la précision déjà acquises. En **Deep Synergy Learning (DSL)**, la situation se gère plus sagement grâce à la **dynamique** auto-organisée qui régit la formation et la déliaison des connexions.

Au fil du temps, la **synergie** entre entités se réévalue en continu. Si un groupe de capteurs cesse de fournir une information fiable, les pondérations associées baissent (faible synergie), déconnectant progressivement l'entité devenue obsolète de la structure ; à l'inverse, si de nouvelles variables ou de nouvelles relations s'avèrent profitables, elles consolident leurs liaisons. Le réseau se remodèle donc de manière **locale**, sans nécessiter la rétropropagation globale ni un réapprentissage exhaustif.

Cette **localité** de la mise à jour, où chaque entité s'occupe prioritairement de ses **liens directs** (et s'ajuste indirectement via les liaisons d'autres entités), confère au **SCN** une forme de **dynamique distribuée**. Les nœuds décident eux-mêmes de renforcer ou d'affaiblir leurs coopérations, suivant l'intérêt (synergie) rencontré, sans orchestration centrale. Le système **s'adapte** ainsi de manière fluide, maintenant une relative **stabilité** pour les groupes demeurant utiles, tout en acceptant

l'insertion ou la réduction de connexions pour refléter l'évolution réelle de la **distribution** des données.

1.5.4.4. Comparaison avec l'Architecture Figée d'un Réseau Profond

Dans la majorité des **réseaux neuronaux** classiques (CNN, RNN, MLP, Transformer), l'entraînement s'effectue via **rétropropagation** sur un **ensemble de données** donné a priori, aboutissant à un **modèle** fini. Lorsque la distribution des données évolue ou qu'une nouvelle tâche survient, on se retrouve souvent contraint de réentraîner ou de fine-tuner l'architecture, avec un **risque** d'oubli des acquis précédents (catastrophic forgetting) et un **coût** en calcul potentiellement élevé. Cette rigidité découle du fait que les **couches** et la **connectivité** du réseau restent invariables. Les structures telles que les **couches convolutionnelles** ou les **mécanismes de self-attention** demeurent fixes, sans mécanisme natif permettant d'**ajouter** ou de **supprimer** dynamiquement des neurones ou des flux de données.

Le **Deep Synergy Learning (DSL)** propose une philosophie différente. La structure du **Synergistic Connection Network** est en **évolution continue**. Les entités \mathcal{E}_i peuvent mettre à jour leurs **représentations** internes \mathbf{x}_i , \mathbf{s}_i et les **liaisons** $\omega_{i,j}$ s'ajustent (ou se dissolvent) selon le niveau de **synergie**. On peut donc introduire une **nouvelle entité** lorsqu'un flux de données inédits émerge, ou en ôter une qui ne sert plus, sans devoir **réentraîner** le système dans son intégralité. L'**apprentissage** n'est pas un processus qui se clôt : le **réseau** conserve un état de plasticité lui permettant de **s'adapter** en continu aux transformations de l'environnement. Cette approche relève d'un véritable **lifelong learning**, affectant aussi bien les **paramètres** que la **topologie** du SCN. C'est précisément cette capacité de reconfiguration qui différencie la souplesse du **DSL** de la rigidité d'un réseau profond figé, et qui évite de longues phases de réapprentissage global ou de coûteux correctifs pour incorporer de nouveaux concepts ou signaux.

1.5.4.5. Exemples d'Application Pratique

Il existe plusieurs scénarios où l'**approche de Deep Synergy Learning (DSL)**, avec son **réseau** auto-organisé et ses liens **adaptatifs**, se révèle particulièrement puissante pour faciliter un **apprentissage continu**.

Dans un **contexte de robotique autonome**, l'environnement et les capteurs évoluent en permanence. La luminosité peut varier, de nouveaux obstacles apparaissent, ou un capteur supplémentaire, comme une caméra ou un lidar, est ajouté. Dans un **SCN**, chaque capteur ou module de traitement est représenté par une **entité**. Lorsqu'un capteur devient peu fiable en raison du bruit ou d'une altération de ses mesures, ses connexions ω s'affaiblissent naturellement, réduisant ainsi son influence. À l'inverse, lorsqu'un nouveau dispositif est introduit, il suffit de l'ajouter en tant que **nouvelle entité** $\mathcal{E}_{\text{nouvelle}}$, et la **synergie** avec les autres modules s'évalue automatiquement. Sans reconstruire l'architecture globale, le **DSL** réorganise les connexions $\{\omega_{i,j}\}$ et **réalloue** l'importance de chaque capteur en fonction de sa **fiabilité** et de sa **pertinence** pour la tâche en cours.

Dans un **système de recommandation évolutif**, l'ajout constant de nouveaux utilisateurs et de nouveaux contenus (films, livres, produits...) constitue un défi pour les approches classiques. Les méthodes traditionnelles, comme le **collaborative filtering**, nécessitent souvent un **réapprentissage** périodique ou un recalcul complet des similarités. Avec un **DSL**, chaque nouvel utilisateur ou nouvel objet est simplement **intégré** au **réseau** en tant qu'**entité** \mathcal{E}_{new} , tandis que la dynamique de **synergie** ajuste automatiquement les connexions $\omega_{\text{new}, \dots}$. Les liaisons avec les utilisateurs ou contenus similaires se renforcent progressivement, créant un **cluster** autour de la nouvelle entité. Le recalcul reste **local**, permettant au **réseau** d'évoluer en continu sans nécessiter un **réentraînement global**.

Dans l'**analyse contextuelle**, comme l'étude de flux de données urbaines ou sociales, la **distribution** des données varie en fonction du temps. Un réseau traditionnel devrait prévoir différents modèles pour gérer ces changements ou recalculer ses poids à intervalles réguliers. Dans un **SCN**, les **clusters** pertinents se forment pour un contexte donné et se réactualisent lorsque la situation évolue. Lors d'un passage du **jour** à la **nuît** ou d'une transition entre **saisons**, certains sous-groupes se dissolvent si la synergie disparaît, tandis que d'autres émergent. Le **DSL** ajuste ainsi les connexions en **temps réel**, permettant de commuter fluidement d'un **cluster** à l'autre sans nécessiter un **réentraînement coûteux**.

1.5.5. Réduction de la Dépendance à la Supervision Humaine

Au-delà de la multi-modalité (1.5.2), de la gestion des données incomplètes (1.5.3) et de l'adaptation continue (1.5.4), un autre avantage essentiel du **Deep Synergy Learning (DSL)** réside dans sa **relative autonomie** vis-à-vis d'une supervision humaine intensive. Les méthodes traditionnelles de l'IA, notamment les réseaux neuronaux profonds, reposent souvent sur l'accès à d'importants **jeux de données labellisés**, et requièrent un entraînement supervisé (ou semi-supervisé) afin d'ajuster des poids internes. L'annotation des données ou la configuration d'un réseau supervisé peut s'avérer **coûteuse** et **chronophage**, rendant difficile l'adaptation à des environnements non balisés. De plus, dans de nombreux contextes réels, la **disponibilité des labels** est limitée, et un apprentissage plus **auto-dirigé** devient une nécessité, notamment face à des volumes massifs de **big data non étiqueté** ou à des environnements inconnus.

Le **DSL**, en favorisant des **relations auto-organisées** entre entités, intègre des mécanismes permettant de réduire la dépendance aux **étiquettes externes**. Plutôt que d'être guidé exclusivement par des labels, il s'**auto-structure** en fonction des synergies détectées, ce qui diminue la nécessité d'une supervision stricte. Cette capacité offre des perspectives intéressantes pour l'**IA autonome** et l'**apprentissage non supervisé**, où l'identification des relations internes entre entités permet de révéler des structures pertinentes sans intervention humaine.

1.5.5.1. L'Auto-Organisation sans Label

Dans les paradigmes d'**apprentissage supervisé**, on dispose en général pour chaque donnée \mathbf{x} d'un **label** y (classe, valeur numérique à prédire), et l'entraînement consiste à **minimiser** une fonction de coût $\mathcal{L}(f_{\theta}(\mathbf{x}), y)$. On **ajuste** alors les poids du réseau neuronal pour améliorer la correspondance entre l'entrée \mathbf{x} et la sortie souhaitée y .

Dans le **Deep Synergy Learning (DSL)**, la dynamique s'appuie sur la **synergie** entre paires d'entités \mathcal{E}_i et \mathcal{E}_j . On définit une quantité $S(\mathcal{E}_i, \mathcal{E}_j)$ qui mesure le **gain** (ou la complémentarité) qu'elles retirent de leur coopération. Cette fonction de synergie peut reposer sur une **distance** ou une **similarité**, sur la **co-information** en théorie de l'information, ou encore sur tout **critère** mesurant l'utilité mutuelle. Dans un **réseau** entièrement non supervisé, aucune annotation externe n'est disponible. La **structure** se forme alors naturellement à travers les **interactions** entre entités, qui s'**agrègent** ou se **séparent** en fonction de leur synergie locale. Un **objectif** global, le cas échéant, peut exister si un signal partiel ou une contrainte externe oriente la cohérence, mais il n'est pas obligatoire. L'auto-organisation résulte de la **conjonction** des ajustements locaux, sans référence à un label y .

Cette logique se reflète dans la **formation** de clusters (section 1.4.3). Les entités se regroupent par **renforcement** de leurs connexions si elles détectent un **gain**, et se dissocient sinon. Il n'est pas nécessaire de leur **attribuer** des étiquettes telles que "classe C1" ou "classe C2". Elles se regroupent **spontanément**, sans supervision, en fonction de la **synergie perçue**, formant ainsi des structures auto-organisées. Le processus aboutit à un **clustering** naturel, que l'on peut ensuite interpréter comme un groupement de similarités ou de dépendances, et qui peut se révéler **faiblement supervisé** si l'on dispose, en parallèle, de quelques **signaux** ou **contraintes** externes. Dans ce cadre, le **DSL** encourage un **apprentissage local** des coopérations, structurant le **Synergistic Connection Network** sans recourir à un label global pour chaque exemple.

1.5.5.2. Quand un Signal de Supervision Existe...

Le **Deep Synergy Learning (DSL)** n'exclut pas l'éventualité de **labels** ou d'**objectifs** supervisés. Au contraire, il est tout à fait envisageable de disposer d'un **objectif** supervisé pour prédire un label y , tout en conservant le mécanisme d'**auto-organisation** qui ajuste les liens $\omega_{i,j}$ en fonction de la synergie. Dans ce scénario, le **DSL** cohabite avec un algorithme de rétropropagation ou d'autre forme de supervision, mais n'est pas entièrement asservi à cet unique signal étiqueté.

Concrètement, on peut imaginer qu'un **objectif** tel que la minimisation d'une fonction de perte $\mathcal{L}(f_\theta(\mathbf{x}), y)$ se déroule en parallèle de la **dynamique** locale guidée par la synergie. Les **liaisons** $\omega_{i,j}(t)$ poursuivent la règle adaptative (section 1.4.5), tandis que le **label** y n'intervient que pour évaluer la cohérence globale ou orienter certains choix de structure. Une partie du réseau peut ainsi se spécialiser dans la tâche supervisée, la **synergie** jouant un rôle déterminant dans l'**organisation** des entités restantes.

Cette architecture présente un **double** bénéfice. D'abord, l'**auto-organisation** ne repose pas exclusivement sur le label y , ce qui permet un fonctionnement **semi-supervisé** ou même **non supervisé**. La majorité des liaisons $\omega_{i,j}$ se forment en fonction de la **coopération** intrinsèque entre entités, tandis que l'objectif supervisé, s'il est présent, n'intervient qu'en **complément**. Ensuite, si le nombre de labels est **limité**, l'apprentissage reste largement alimenté par la **synergie** interne, évitant la nécessité de disposer d'un jeu de données richement annoté. Le **réseau** peut ainsi, en un premier temps, **structurer** ses entités par pure auto-organisation, puis tirer parti de quelques **labels** additionnels pour affiner la prédiction d'une variable ou la détection d'anomalies.

L'important est que les **labels**, quand ils existent, ne contrôlent pas toutes les **liaisons** du réseau. Seule une fraction des entités ou des connexions peut être orientée par la performance supervisée, tandis que les **clusters** naissent en grande partie de la **synergie** (sections 1.4.3 et 1.4.4). Le réseau se montre donc apte à **apprendre** et **s'adapter** via son auto-organisation, puis à **exploiter** un signal supervisé pour ajuster un module final ou un sous-ensemble d'entités, évitant ainsi la contrainte d'une rétropropagation exhaustive sur toute l'architecture.

1.5.5.3. Exemple : Découverte Spontanée de Catégories

Pour illustrer la manière dont un **Deep Synergy Learning (DSL)** peut structurer les données **sans** s'appuyer sur des labels explicites, on peut considérer un ensemble de **documents** non annotés. Chaque document est alors représenté par un **embedding** ou un **bag-of-words**. Dans le **Synergistic Connection Network**, on introduit autant d'entités \mathcal{E}_{doc} que de documents, en laissant le mécanisme de **synergie** déterminer comment ces entités se relient.

Les entités commencent par **évaluer** leur proximité ou leur complémentarité. Par exemple, des vecteurs \mathbf{x}_{doc} représentant des documents partageant des thèmes communs, un vocabulaire similaire ou des distributions de topics proches peuvent manifester une **synergie** significative. Les pondérations $\omega_{i,j}$ entre entités se renforcent si la **similarité** est jugée élevée, ou si la co-information (section 1.4.4) révèle un **gain** dans la mise en commun de leurs contenus. Progressivement, un ou plusieurs **clusters** apparaissent, chaque cluster rassemblant des documents sur des sujets proches, sans qu'on ait besoin de spécifier de labels ou de classes. Ainsi, certains regroupements peuvent porter sur la **politique**, d'autres sur le **sport**, d'autres encore sur la **santé**, mais rien n'impose que le DSL les nomme ou les identifie explicitement.

Si, plus tard, un **expert** fournit quelques **labels** indiquant, par exemple, qu'un certain cluster correspond au domaine sportif, on n'a pas besoin d'annoter l'**intégralité** des documents. Le **DSL** s'étant déjà **auto-organisé** en clusters cohérents, il suffit d'**attribuer** une étiquette à la communauté d'entités concernée une fois la structure formée. Ce procédé montre comment un système **non supervisé** ou **faiblement supervisé** peut accomplir une **structuration** fine des données, sur laquelle un label partiel ne fait qu'ajouter un niveau de **description** plus explicite. Les sections antérieures (1.4.3, 1.5.5.1 et 1.5.5.2) soulignent ainsi la flexibilité du DSL pour fonctionner dans des contextes sans labels, tout en permettant une intégration de signaux supervisés quand ils existent.

1.5.5.4. Vers une IA plus Autonome

Les **interactions indirectes** (discutées en section 1.4.6) font en sorte qu'une entité peut être **influencée** par une autre, même en l'absence de lien direct, via des chemins de coopération au sein du **Synergistic Connection Network**. Ce mécanisme enrichit le potentiel d'**auto-découverte** de catégories ou de concepts de manière **non supervisée**, puisque l'information peut circuler librement à travers divers **chemins**, et les **clusters** émergent en valorisant l'intégralité des signaux disponibles.

Dans une perspective de **IA** plus générale, appelée à affronter des **domaines inconnus** et des **contextes** évolutifs, la **réduction** du besoin de supervision revêt un caractère décisif. Le **Deep Synergy Learning (DSL)**, en misant sur un **apprentissage local** et **continu**, se montre apte à fonctionner avec des annotations partielles ou inexistantes, tout en conservant la capacité de **structurer** et **organiser** les entités. Il s'agit d'une étape vers une IA plus autonome, qui, sans nécessité d'un étiquetage exhaustif ou de guides fortement supervisés, s'adapte aux environnements **faiblement annotés** (voire non annotés) et continue à explorer la structure interne des données pour faire émerger **concepts** et **catégories** spontanément.

1.5.5.5. Limites et Solutions Partielles

Le **Deep Synergy Learning (DSL)** ne s'appuie pas massivement sur les **labels**, ce qui lui confère une grande liberté dans la formation des **clusters** et dans la découverte de structures internes. Néanmoins, un **petit** signal externe, issu par exemple d'une **semi-supervision**, peut être requis lorsque l'on souhaite orienter la **dynamique** vers des objectifs spécifiques. Sans cette impulsion, le réseau se contente de former des groupes qui font sens pour les données elles-mêmes, sans garantir qu'ils correspondent aux besoins d'une application particulière. Il peut arriver que les clusters, très cohérents sous un angle statistique, ne concordent pas avec la segmentation ou les catégories souhaitées par l'utilisateur.

L'**auto-organisation** peut en effet aboutir à des **partitions** intrinsèquement pertinentes, mais qui ne recouvrent pas exactement les **objectifs** fixés en pratique. Pour rectifier ce phénomène, il est possible d'introduire un **minimum** de **contraintes** ou de **labels** annotés, lesquels jouent le rôle de **pénalités** ou de **feedback** dans la mise à jour des liaisons $\{\omega_{i,j}(t)\}$. De la sorte, le DSL ne se limite plus à la seule synergie intrinsèque, mais tient également compte de la finalité appliquée (une tâche métier, une classification imposée), ce qui oriente plus étroitement la **constitution** des clusters.

La **complexité** du calcul de la synergie, surtout pour des données brutes très dimensionnelles, constitue un autre défi. Dans certains cas, on envisage des mesures plus poussées telles que la **co-information** ou la **Partial Information Decomposition** (sections 1.4.4 et 1.4.7), dont l'estimation peut s'avérer lourde. Des **estimateurs** non paramétriques requièrent souvent un échantillonnage important et des ressources conséquentes. Il est alors crucial de se doter de **stratégies** d'approximation ou d'**échantillonnage** (sampling parcimonieux, heuristiques de calcul) pour maintenir la faisabilité sur des volumes de données massifs (big data). Cette solution, certes partielle, demeure incontournable si l'on veut étendre le **DSL** à des scénarios industriels exigeants, tout en préservant un **coût** de calcul acceptable.

1.5.6. Création de Représentations Riches et plus Interprétables

En plus de la **réduction de la dépendance à la supervision** (1.5.5), le **Deep Synergy Learning (DSL)** présente un avantage notable pour la **création de représentations** à la fois **riches** et **interprétables**. Contrairement aux réseaux neuronaux profonds classiques, qui produisent souvent des représentations “boîte noire” difficilement explicables, le DSL mise sur l'**auto-organisation** et la **coopération** entre entités d'information, ce qui peut aboutir à des structures internes (clusters,

macro-clusters, liens synergiques) plus compréhensibles pour un humain. Cette section (1.5.6) met en évidence :

- **Pourquoi** la synergie favorise l'émergence de **représentations complexes**,
- Comment le principe de **clusters** ou de **macro-clusters** rend l'organisation plus **lisible**,
- En quoi le DSL peut faciliter l'**interprétabilité** par rapport aux architectures profondes traditionnelles.

1.5.6.1. Logique “par Entités” plutôt que “en Couches Opaques”

Dans les **réseaux neuronaux profonds** (CNN, RNN, Transformers...), la **représentation interne** repose sur une succession de **couches** appliquant des transformations non linéaires. Malgré certaines avancées en matière d'**interprétabilité**, comme la visualisation des **filtres** dans les CNN ou l'analyse des **matrices d'attention** dans les Transformers, comprendre comment le réseau parvient à une décision globale reste un défi majeur.

L'une des principales limitations réside dans l'**absence de structuration explicite** des poids. Ceux-ci ne sont pas naturellement regroupés sous forme de **blocs lisibles**, ce qui complique l'identification de sous-ensembles de paramètres correspondant à des **concepts** distincts. À l'exception d'un neurone ou d'une couche spécifique, il est souvent difficile de relier une partie du réseau à une interprétation claire de son rôle.

Les méthodes d'**interprétabilité** existantes, comme **Grad-CAM** ou **LIME**, tentent de pallier cette opacité en mettant en évidence l'influence de certaines **caractéristiques locales** (pixels, dimensions spécifiques des embeddings). Toutefois, ces techniques reposent généralement sur des **heuristiques** qui, bien qu'utiles pour identifier des tendances, ne garantissent pas une **compréhension globale** de la logique interne du réseau.

Le **Deep Synergy Learning (DSL)** adopte une approche fondamentalement différente en organisant l'information autour d'**entités**, qui constituent les **nœuds** du **Synergistic Connection Network (SCN)**. Contrairement aux réseaux neuronaux traditionnels où les poids sont répartis sans structuration explicite, le **DSL** permet de **modéliser** chaque entité comme un **flux sensoriel** (image, son, texte) ou un **descripteur abstrait** (embedding, concept partiel).

Les **liaisons** entre entités sont définies par la **synergie** détectée, c'est-à-dire la **plus-value mutuelle** qu'elles retirent de leur coopération. Cela transforme la représentation en un **graphe modulaire**, où l'on peut **visualiser** les **interactions clés** et observer comment différentes entités collaborent. Contrairement aux architectures classiques, où l'information circule de manière séquentielle à travers des **couches opaques**, le **DSL** permet une **organisation plus flexible**, où les **clusters** émergent de manière dynamique en fonction des **relations synergiques**.

Cette structuration améliore la **modularité** et la **lisibilité** du réseau. Par exemple, un **cluster** peut regrouper la **détection de visages** (vision), l'**audio de voix** (son) et un **module de reconnaissance textuelle** (analyse des mots-clés parlés), formant ainsi une unité cohérente qui associe plusieurs modalités de manière fluide. Contrairement à l'**empilement rigide** des couches dans les

architectures classiques, cette approche permet une **segmentation plus fine** des interactions et une **interprétabilité accrue**.

Grâce à cette conception **basée sur les entités**, le SCN offre une représentation plus **transparente**. Il devient possible d'**identifier précisément** les **entités en présence**, de connaître leurs **rôles spécifiques**, et de comprendre comment elles s'**agencent** via les **pondérations de synergie**. Le DSL apporte ainsi un degré supplémentaire d'**explicabilité**, en offrant une **cartographie des flux d'information**, bien plus lisible que l'accumulation complexe et parfois opaque de poids dans un réseau neuronal profond traditionnel.

1.5.6.2. Clusters et Macro-Clusters comme Briques de Sens

Dans le **Deep Synergy Learning (DSL)**, la formation d'un **cluster** d'entités traduit souvent l'émergence d'un **concept** ou d'un **thème** sous-jacent aux données. Par exemple, un **cluster** multimodal composé d'entités liées à l'image, à l'audio et au texte pourrait se spécialiser dans la détection d'un **événement conférence**. L'entité image capterait une scène où un orateur fait face à un public, l'entité audio identifierait des bruits de voix ou des applaudissements, et l'entité texte reconnaîtrait des termes comme "bienvenue" ou "question". Cette **cohérence** ne se réduit pas à un simple ajustement de poids numériques. Elle permet de comprendre quelles entités s'agrègent et pour quelle raison, grâce à l'évaluation locale de la **synergie**, ce qui améliore la **lisibilité** du modèle.

D'un point de vue **explicatif**, un **cluster** dans le **Synergistic Connection Network (SCN)** ne se limite pas à un agrégat opaque de neurones, comme dans un réseau profond traditionnel. Il s'apparente plutôt à un **sous-graphe** composé d'entités distinctes, qu'il s'agisse de capteurs visuels, de flux audio, de modules textuels ou encore de représentations symboliques. Les **liaisons synergiques** entre ces entités reflètent la force de leur coopération et permettent d'analyser la distribution de la **co-information** ou de la **similarité** entre elles, rendant le processus plus transparent.

Lorsque plusieurs **clusters** se renforcent mutuellement, ils peuvent se **fusionner** pour former un **macro-cluster**. Cette dynamique révèle que des groupes d'entités s'organisent en structures plus larges et interconnectées. Dans un **contexte e-commerce**, un macro-cluster pourrait ainsi regrouper plusieurs **catégories de produits**, des **groupes d'utilisateurs** et des **attributs de profil**, mettant en évidence un **sous-réseau dynamique** où la synergie favorise des interactions complexes et évolutives.

Ces **briques** constituées de **clusters** et de **macro-clusters** forment des **modules interprétables** que l'on peut analyser et commenter. Chaque entité, en raison de son association avec un **flux de données** ou un **vecteur caractéristique**, possède un **ancrage explicite** qui clarifie son rôle au sein du réseau. Les **liaisons synergiques** révèlent la proportion de **coopération interne**, tandis que le DSL segmente naturellement le graphe global en plusieurs **sous-ensembles** structurés autour de **concepts** ou de **thématiques**.

Cette granularité confère au modèle une **modularité accrue**, facilitant une analyse plus fine de son organisation. L'approche permet d'examiner directement les **relations** et les **interactions** au sein du réseau, offrant ainsi une compréhension plus accessible que l'interprétation classique des poids dans une **couche neuronale abstraite**. Grâce à cette structuration, il devient possible d'identifier

des justifications tangibles sur le fonctionnement du système et d'expliquer plus clairement les décisions qu'il génère.

1.5.6.3. Représentations Multiples et Non Linéaires

Le **Deep Synergy Learning (DSL)** ne se borne pas à des représentations linéaires ou statiques. Au contraire, chaque **entité** \mathcal{E}_i peut conserver une **représentation interne** $\mathbf{x}_i(t)$ ou $\mathbf{s}_i(t)$ qui évolue au fil du temps, ou être associée à un petit réseau neuronal local (par exemple un autoencodeur) calculant

$$\mathbf{x}_i(t + 1) = F_i(\mathbf{x}_i(t), \{\omega_{i,k}(t)\}, \dots).$$

La **non-linéarité** inhérente à ce type de module (RNN local, MLP, autoencodeur) élargit considérablement l'espace des **représentations** envisageables, comparativement à un simple cadre vectoriel fixe. Les entités peuvent ainsi affiner leurs **features** internes ou leurs **états** \mathbf{s}_i , tout en mettant à jour leurs **liaisons** $\omega_{i,j}(t)$ d'après la **synergie** nouvellement constatée. Cette approche ne repose pas sur un pipeline rigide. Chaque entité, conçue comme un **bloc fonctionnel autonome**, conserve sa propre dynamique d'évolution. Pendant ce temps, la **synergie** entre les entités se réévalue localement en fonction des interactions et des changements contextuels. Cette flexibilité permet au système d'adapter continuellement ses connexions, sans nécessiter une reconfiguration globale imposée.

Si deux entités en viennent à **coïncider** fortement (en partageant des représentations très similaires ou très complémentaires), leurs pondérations $\omega_{i,j}$ peuvent atteindre un niveau **saturé**, conduisant potentiellement à une **fusion** (au sens de la fusion de clusters, section 1.4.3). Inversement, d'autres entités peuvent demeurer **indépendantes**, préférant garder leurs liaisons minimales. Le résultat global est une **représentation** dite "multi-foyers", où plusieurs **clusters** se spécialisent dans des sous-ensembles de la tâche ou des données. Chacun possède son propre noyau fonctionnel, et ils coopèrent occasionnellement si la **synergie** l'exige. Cette organisation **modulaire** tranche avec un réseau profond monolithique dans lequel toutes les informations finissent entremêlées au sein de couches successives. Elle favorise un **équilibre** entre la **richesse** (via la non-linéarité) et la **lisibilité** (chaque entité reste limitée à un bloc fonctionnel clairement cerné), tout en laissant l'auto-organisation diriger la consolidation ou la séparation des diverses composantes.

1.5.6.4. Comparaison avec l'Explicabilité dans les Réseaux Profonds

Dans les approches classiques d'**explainability** appliquées aux réseaux neuronaux profonds (CNN, Transformers, etc.), il est habituel de recourir à des méthodes "post-hoc", comme la visualisation d'**activation maps**, l'explication locale (LIME, SHAP) ou l'examen des **attention maps** dans un Transformer. Ces techniques fournissent un aperçu de la raison pour laquelle le réseau met en avant tel pixel, telle dimension ou tel token, mais elles ne suppriment pas pour autant le caractère opaque de la structure interne. Un CNN reste un empilement de couches, au sein desquelles la signification des poids demeure largement cachée, et un Transformer recèle une superposition de blocs d'attention dont la lisibilité globale n'est pas toujours aisée.

Le **Deep Synergy Learning (DSL)** propose une **explication intrinsèque** plus claire en structurant l'information sous forme d'entités reliées par des **liaisons synergiques** $\omega_{i,j}$. Ces connexions traduisent la coopération effective entre les entités, ce qui permet d'obtenir un **graphe lisible** où apparaissent naturellement des **clusters** \mathcal{C} et des **macro-clusters** représentant des groupes d'entités en interaction. Chaque entité conserve une **identité distincte**, qu'elle soit liée à une modalité spécifique, à un bloc fonctionnel ou à un concept partiel.

L'utilisateur peut examiner un **cluster** \mathcal{C} pour identifier les entités qui le composent, comprendre leurs caractéristiques et analyser la distribution des pondérations $\omega_{i,j}$. Cette structure permet d'interpréter la **logique d'assemblage** des sous-groupes formés. Par exemple, un sous-graphe associant la détection de panneaux de signalisation, l'enregistrement de la vitesse et le suivi de la trajectoire montre clairement la raison d'être de cette coopération à travers la synergie mesurée par les valeurs de $\omega_{i,j}$.

Cette transparence contraste avec les réseaux neuronaux profonds traditionnels, où les ensembles de neurones internes restent difficiles à interpréter sans recourir à des méthodes explicatives externes. Le **DSL** rend ces connexions **explicites et modulables**, offrant une vision claire des interactions et des coopérations à travers les **pondérations synergiques** $\omega_{i,j}$ du **Synergistic Connection Network**.

1.5.6.5. Exemple d'Application : Représentation Sémantique Évolutive

Dans un **agent conversationnel**, plusieurs **entités** interagissent en fonction des volets de l'échange. Une entité $\mathcal{E}_{\text{linguistique}}$ se charge de la compréhension du langage, tandis qu'une entité $\mathcal{E}_{\text{contexte}}$ gère les informations externes comme l'heure, le lieu ou l'historique de conversation. En parallèle, une entité $\mathcal{E}_{\text{profil utilisateur}}$ reflète les préférences et les habitudes du locuteur.

Au fil des interactions, les synergies entre ces entités évoluent : lorsque $\mathcal{E}_{\text{profil}}$ apporte des indices alignés avec $\mathcal{E}_{\text{linguistique}}$, les pondérations $\omega_{i,j}$ entre ces composants augmentent. Le **DSL** favorise alors la formation d'un **cluster** $\mathcal{C} = \{\mathcal{E}_{\text{ling}}, \mathcal{E}_{\text{contexte}}, \mathcal{E}_{\text{profil}}\}$, où la coopération entre ces entités devient explicite à travers l'évolution des liens $\omega_{i,j}$.

Ce processus conduit à l'émergence d'un **sous-modèle** sémantique organisé autour des thèmes, des intentions et des préférences récurrentes du locuteur. En analysant la structure du **Synergistic Connection Network**, il devient possible d'identifier les **topics** dominants et de comprendre comment ces synergies ont émergé au sein du réseau.

Dans un autre registre, l'**analyse d'images documentaires** bénéficie également du **DSL** grâce à la structuration des informations en **entités** interconnectées. Certaines entités $\mathcal{E}_{\text{patch}}$ décrivent des patches ou segments visuels, représentant des éléments d'architecture ou des styles picturaux. D'autres entités regroupent des **mots-clés** extraits des légendes ou des métadonnées associées aux images, tandis qu'un troisième type d'entités capture des **concepts** plus abstraits liés à une époque, un courant artistique ou un auteur.

Lorsque ces entités présentent une synergie significative, un **cluster** émerge spontanément. Par exemple, un patch visuel $\mathcal{E}_{\text{patch}}$ contenant une structure architecturale gothique peut se lier à un

ensemble de mots-clés \mathcal{E}_{mot} décrivant des éléments médiévaux, ainsi qu'à une entité conceptuelle $\mathcal{E}_{\text{concept}}$ indiquant la période historique concernée. L'auto-organisation du **Synergistic Connection Network** révèle alors une **co-information** élevée : les motifs visuels, les mots-clés et l'ancrage conceptuel convergent vers une interprétation unifiée.

Cette **coopération** entre entités confère une **lisibilité accrue** : sans imposer de règles prédéfinies, le réseau identifie naturellement un regroupement signifiant « document archéologique médiéval ». Le **DSL** agit ainsi comme un **pilier** sémantique dynamique, renforçant les connexions pertinentes lorsqu'un bénéfice mutuel est détecté et permettant une **représentation** riche, évolutive et modulable.

1.5.7. Intégration de Dimensions Symboliques ou Cognitives

Les sections précédentes (1.5.1 à 1.5.6) ont souligné diverses forces du **Deep Synergy Learning (DSL)**, telles que la **capacité à gérer la multi-modalité**, à **s'adapter** en continu, ou à **créer** des représentations plus **lisibles**. Un autre point capital, souvent abordé dans l'IA contemporaine, est la possibilité de **mélanger** des composantes **sub-symboliques** (issues des méthodes connexionnistes) et des composantes **symboliques** (notions logiques, sémantiques, règles expertes). Alors que les réseaux neuronaux profonds "classiques" ont généralement peiné à intégrer ces dimensions symboliques, le **DSL**, par sa structure en entités et liens synergiques, peut être un **cadre** propice pour unir symbolique et sub-symbolique.

Cette section (1.5.7) explique :

- Comment le DSL autorise l'**introduction** d'entités à **portée symbolique** (règles, concepts, modèles cognitifs),
- Pourquoi cette intégration est **plus fluide** que dans un réseau hiérarchique figé,
- Les **applications** possibles (raisonnement, inférence, cognition proche de l'humain),
- Les **défis** mathématiques associés à la cohabitation de ces deux modes de représentation.

1.5.7.1. Aperçu des Approches Symboliques et Connexionnistes en IA

L'**intelligence artificielle** repose historiquement sur deux grandes approches aux fondements distincts. L'**IA symbolique** s'appuie sur des **règles logiques** et des **moteurs d'inférence**, permettant de manipuler des symboles et d'expliquer chaque décision par une suite de règles compréhensibles. Cette approche garantit une **transparence** et une **cohérence** dans le raisonnement, ce qui facilite son interprétabilité. Toutefois, elle se heurte à des limites face aux **données bruitées** et aux environnements évolutifs, car elle peine à s'adapter et à extraire des **motifs complexes** directement depuis des observations non structurées.

À l'opposé, l'**approche connexionniste** repose sur l'apprentissage automatique et l'ajustement dynamique des **paramètres internes** des **réseaux neuronaux**. Ces modèles sont capables de détecter des régularités dans de vastes ensembles de données et d'apprendre sans qu'un programme

explicite ne dicte leurs règles. Cette plasticité leur confère une grande efficacité dans des tâches comme la reconnaissance d'images ou la compréhension du langage naturel. Toutefois, cette approche souffre d'un problème majeur. Les décisions émergent d'un **réseau dense de pondérations**, rendant leur fonctionnement difficile à interpréter. Contrairement à l'IA symbolique, ces modèles n'intègrent pas de **structures logiques explicites** et ne possèdent pas de mécanismes de **raisonnement formel**, ce qui complique leur fiabilité et leur explicabilité.

Face à ces enjeux, le **Deep Synergy Learning (DSL)** propose une approche qui intègre les atouts des deux paradigmes. Plutôt que d'opposer **symbolisme** et **connexionnisme**, le **Synergistic Connection Network (SCN)** combine des **concepts symboliques** avec des **représentations neuronales** issues de l'apprentissage automatique. Cette hybridation permet de conserver la **plasticité adaptative** des modèles neuronaux tout en structurant mieux les connaissances acquises. L'**explicabilité** est renforcée par l'analyse des **liaisons synergiques** $\omega_{i,j}$, qui offrent une vision plus modulaire et interprétable de l'apprentissage.

Cette approche ouvre la voie à une **intelligence artificielle plus robuste**. Le réseau peut **évoluer continuellement**, **justifier ses choix** et **intégrer des connaissances structurées**, tout en restant capable de s'adapter aux nouvelles données et aux variations du contexte.

1.5.7.2. Principes d'une Approche Hybride Symbolique–Connexionniste

Dans la continuité des sections précédentes, et en particulier au regard des principes généraux du **Deep Synergy Learning (DSL)**, on constate que les approches **symboliques** et **sub-symboliques** apparaissent de plus en plus **complémentaires**. Les récentes avancées en **IA** mettent en évidence cette complémentarité. D'un côté, les approches sub-symboliques, comme les **réseaux neuronaux classiques**, le **DSL** et d'autres paradigmes similaires, offrent une **puissance d'apprentissage** notable pour traiter de grandes quantités de données et extraire des **caractéristiques complexes** dans des espaces de haute dimension. De l'autre, les méthodes symboliques (appuyées sur des **logiques**, des **règles** et des **ontologies**) assurent une **rigueur** et une **explicabilité** supérieures, notamment lorsqu'il convient de formuler des raisonnements structurés ou de décrire des **connaissances** de manière **interprétable**.

La question est donc de **concevoir** une architecture **hybride** qui mobilise simultanément la **puissance d'apprentissage** sub-symbolique et la **cohérence symbolique**. Dans le cadre d'un **Synergistic Connection Network (SCN)** tel que défini dans les sections précédentes, il est possible d'**introduire** des entités de différents types. Il est ainsi envisageable de faire cohabiter des **entités** $\mathcal{E}_{\text{symb}}$ représentant des **concepts logiques**, des **règles formelles**, ou des **faits** structurés, et des **entités** $\mathcal{E}_{\text{data}}$ se fondant sur des **vecteurs sub-symboliques** $\mathbf{x}_i \in \mathbb{R}^d$. On peut imaginer que la représentation symbolique contienne des énoncés tels que « Un véhicule se déplace sur roues » ou « Si la température excède $100\text{ }^{\circ}\text{C}$, on suspecte une surchauffe », tandis que la partie sub-symbolique gère des **flux de données** (images, séries temporelles, signaux sensoriels) ou des **features** issues de réseaux de neurones.

Dans cette optique, on peut noter $\mathcal{E}_{\text{symb}}$ pour une **entité symbolique** et $\mathcal{E}_{\text{data}}$ pour une **entité sub-symbolique**. La **pondération** $\omega_{\text{symb,data}}(t)$ s'adapte alors par la règle classique du **DSL** :

$$\omega_{\text{symb,data}}(t + 1) = \omega_{\text{symb,data}}(t) + \eta [S(\mathcal{E}_{\text{symb}}, \mathcal{E}_{\text{data}}) - \tau \omega_{\text{symb,data}}(t)],$$

où la **synergie** $S(\mathcal{E}_{\text{symb}}, \mathcal{E}_{\text{data}})$ traduit à quel point la **règle** ou le **concept** véhiculé par $\mathcal{E}_{\text{symb}}$ s'applique utilement au **contenu** sub-symbolique porté par $\mathcal{E}_{\text{data}}$. La valeur de S peut tenir compte du **taux de succès** lors de la confrontation aux données, ou évaluer la **compatibilité** sémantique entre la règle symbolique et les caractéristiques identifiées dans le **flux sub-symbolique**.

On peut approfondir cette **coopération** à l'aide de **formules** décrivant la mise en correspondance entre un **concept** symbolique et des **features** extraites. Par exemple, on peut écrire :

$$S(\mathcal{E}_{\text{symb}}, \mathcal{E}_{\text{data}}) = \text{score}(\text{matching}(\text{rules}(\mathcal{E}_{\text{symb}}), \mathbf{x}_{\text{data}})),$$

où $\text{rules}(\mathcal{E}_{\text{symb}})$ dénote l'ensemble des règles ou faits associés à la représentation symbolique, et \mathbf{x}_{data} la **représentation vectorielle** de l'entité sub-symbolique. La fonction matching évalue la cohérence entre la **logique symbolique** et les **observations** sub-symboliques, et score traduit le **gain** ou la **conformité** obtenu.

Dans cette approche **hybride**, il y a un **avantage** notable. Le réseau **auto-organisé** du DSL peut **identifier** quelles **règles** ou **concepts** sont les plus pertinents pour tel **type de données**, sans nécessiter une supervision massive. Les entités symboliques, en s'activant sur certaines conditions, peuvent **expliquer** plus aisément les **décisions** ou **groupements** effectués. La **transparence** en est renforcée. Si la **pondération** $\omega_{\text{symb,data}}$ augmente, cela signifie que la règle symbolique s'applique fréquemment et avec succès à la modalité de données concernée.

La **limite** principale repose sur la **cohérence** nécessaire entre la **logique** et les **données**. L'entité symbolique doit être suffisamment bien définie et adaptée à la réalité sub-symbolique ; en cas de décalage, la pondération associée risque de rester faible, rendant la règle peu intégrée. Sur le plan de l'implémentation, il faut également gérer la **complexité** liée à l'activation symbolique, potentiellement coûteuse s'il faut évaluer de nombreuses règles sur un flux volumineux de données.

Dans ce **paradigme**, la **puissance** du DSL est préservée. Le réseau se reconfigure en continu, ajustant les liens entre entités symboliques et sub-symboliques en fonction de la **synergie** détectée. En parallèle, les entités symboliques aident à **structurer** le raisonnement et à **rendre compte** de la logique sous-jacente, contribuant à l'**explicabilité** du système final. C'est cette double souplesse qui, dans la ligne de ce chapitre (section 1.5.7.2), semble prometteuse pour de nombreuses applications, allant de la **robotique cognitive** à l'**intégration** de vastes **bases de connaissances** et de **flux sensoriels**.

1.5.7.3. Mécanismes d'Auto-Organisation Intégrant la Logique

Dans la continuité de l'approche **hybride** symbolique–connexionniste présentée en section 1.5.7.2, il est envisageable d'introduire au sein d'un **Synergistic Connection Network (SCN)** des **règles symboliques** interagissant avec des **entités** sub-symboliques. On peut par exemple considérer une **règle** R stipulant :

Si la caméra détecte une forme circulaire, alors suspecter un panneau de signalisation.

Parallèlement, on introduit une **entité sub-symbolique** \mathcal{E}_{cam} chargée d’analyser des données visuelles. Cette entité exploite typiquement un **modèle** neuronique local pour reconnaître ou segmenter des formes dans un **flux** d’images.

La **pondération** $\omega_{R,\text{cam}}$ reliant la règle R à l’entité \mathcal{E}_{cam} suit la dynamique du **DSL**. On peut ainsi l’écrire :

$$\omega_{R,\text{cam}}(t + 1) = \omega_{R,\text{cam}}(t) + \eta [S(R, \mathcal{E}_{\text{cam}}) - \tau \omega_{R,\text{cam}}(t)],$$

où

$S(R, \mathcal{E}_{\text{cam}})$ mesure la **synergie** entre la **règle symbolique** et la **détection sub-symbolique** opérée par \mathcal{E}_{cam} . Concrètement, si la **forme circulaire** repérée par la caméra coïncide fréquemment avec la proposition « panneau de signalisation », la quantité $S(R, \mathcal{E}_{\text{cam}})$ prend une **valeur positive**. Le terme $\eta [S - \tau \omega]$ reste alors **positif**, ce qui **renforce** la liaison $\omega_{R,\text{cam}}$. Progressivement, on voit s’**organiser** un **cluster** local $\{R, \mathcal{E}_{\text{cam}}\}$, soulignant la coopération entre la **règle symbolique** et la **détection** d’images.

Lorsque survient une **deuxième** entité, $\mathcal{E}_{\text{limit}}$, qui incarne un **concept** tel que « panneau de limitation de vitesse », la **coopération** peut encore se prolonger. Si un **module** \mathcal{E}_{ocr} découvre régulièrement le texte “30 km/h” à l’intérieur des formes circulaires détectées, la synergie $S(\mathcal{E}_{\text{ocr}}, \mathcal{E}_{\text{limit}})$ peut s’avérer **élevée**. La **pondération** $\omega_{\text{ocr},\text{limit}}$ croît alors, impliquant l’émergence d’un **sous-réseau** plus vaste :

$$\{R, \mathcal{E}_{\text{cam}}, \mathcal{E}_{\text{ocr}}, \mathcal{E}_{\text{limit}}\}.$$

À l’intérieur de ce **cluster**, les entités **symboliques** (par exemple R et $\mathcal{E}_{\text{limit}}$) formulent des règles logiques ou des connaissances à caractère explicite. Les **entités** \mathcal{E}_{cam} et \mathcal{E}_{ocr} se chargent au contraire d’un **traitement** sub-symbolique (détection de formes, reconnaissance du texte “30”). La **dynamique** des pondérations, selon la règle linéaire ou d’autres variantes (voir section 1.4.5), veille à **consolider** cet ensemble si les corrélations s’avèrent régulières et profitables.

Dans cette **topologie**, le **raisonnement** demeure essentiellement **distribué** et local. Plutôt qu’une structure hiérarchique fixe, l’**auto-organisation** en **micro-réseaux** permet aux règles et aux modules neuronaux de s’**associer** dynamiquement dès que leur **synergie** devient significative. Les **avantages** de cette intégration résident dans la capacité à **exploiter** la puissance de l’apprentissage sub-symbolique pour reconnaître des patterns visuels complexes, tout en s’appuyant sur la **précision** et la **transparence** qu’offrent les entités symboliques. De surcroît, on peut ainsi relier de multiples règles, chacune s’appliquant à des **entités** sub-symboliques différentes, ce qui permet d’atteindre une **modularité** plus élaborée à l’intérieur du **SCN**.

Les **limites** et **inconvenients** tiennent essentiellement au besoin de concevoir des règles symboliques suffisamment précises pour bien refléter la réalité observée par les flux sub-symboliques. S’il existe un décalage important ou un **bruit** excessif dans la détection, la synergie peut rester faible et empêcher la formation d’un cluster stable. Un autre **défi** concerne la **scalabilité**. Lorsqu’un nombre trop élevé de règles ou d’entités sub-symboliques coexistent, leur mise en correspondance devient coûteuse en calcul. Il est alors nécessaire d’adopter des stratégies de **sparsification** ou de **pruning** pour optimiser la dynamique du réseau, comme détaillé dans la section 1.4.5.

Dans l'ensemble, cette **intégration** de la **logique** dans le **DSL** illustre la possibilité de **fusionner** la robustesse neuronale ou statistique et la **structuration** symbolique. L'**auto-organisation** garantit que seules les entités **règle-module** qui partagent un gain effectif conservent un lien fort, aboutissant à une **cohérence** émergente entre les **concepts** logiques et les **vecteurs** de features dans le **Synergistic Connection Network**.

1.5.7.4. Avantages pour la Cognition et l'Explicabilité

Dans la continuité des sections précédentes, le **Deep Synergy Learning (DSL)** se présente comme un **pont** unifiant deux grandes approches du traitement de l'information. D'une part, la **dimension sub-symbolique** s'attache aux aspects de **perception** et de **features** extraits directement des données (images, sons, textes bruts), héritant de la puissance d'**apprentissage** de type connexionniste. D'autre part, la **dimension symbolique** structure des **concepts**, des **règles** et des **faits** logiques qui ne seraient pas nécessairement détectés par une simple analyse statistique. Par exemple, l'affirmation « un véhicule possède quatre roues, un moteur et est conçu pour se déplacer sur la route » représente une **connaissance explicite** que le réseau peut exploiter pour guider son apprentissage et renforcer certaines associations synergiques.

Lorsqu'un **cluster** (au sens de la section 1.4.3) associe une **règle** formalisée à des **entités perceptives** ancrées dans des vecteurs sub-symboliques, on obtient une **forme d'explication** plus directe de la décision. En effet, plutôt que de faire référence à un poids numérique situé dans une couche profonde du réseau, on peut pointer vers le **signal x** (issu de la représentation sub-symbolique) et la **règle R** (exprimée symboliquement) pour justifier la conclusion. Cette **transparence** s'apparente à une forme d'**explicabilité** (souvent recherchée en IA), où l'on peut réellement indiquer quelles **connaissances** et quelles **observations** ont motivé l'activation finale.

Pour formaliser mathématiquement cette idée, on peut écrire que la **pondération** $\omega_{R,data}(t)$ reliant la règle symbolique \mathcal{E}_R à l'entité sub-symbolique \mathcal{E}_{data} évolue selon :

$$\omega_{R,data}(t+1) = \omega_{R,data}(t) + \eta [S(\mathcal{E}_R, \mathcal{E}_{data}) - \tau \omega_{R,data}(t)],$$

où $S(\mathcal{E}_R, \mathcal{E}_{data})$ mesure l'**adéquation** entre la règle symbolique et les **caractéristiques** perçues dans les données. Une fois cette liaison suffisamment renforcée, on peut interpréter la **décision** ou la **catégorie** apprise en mentionnant explicitement la règle R et le segment de données x qui la **valide**.

Je vais reformuler strictement ce que tu as donné, sans ajouter d'interprétation ou d'explication supplémentaire :

Ce mécanisme illustre la notion de **cognition synergique**. Les **entités** symboliques **codent** la **sémantique** et les **règles** explicites, tandis que les **entités** sub-symboliques **apprennent** à partir des données brutes, en extrayant des **patterns** ou des **features** utiles. Dans le **Synergistic Connection Network (SCN)**, la **dynamique** de mise à jour des **pondérations** relie directement ces deux pôles. On aboutit alors à une forme d'**intégration** qui se rapproche de la **cognition humaine**, où des **concepts** explicitement formulés cohabitent avec des **associations perceptives** acquises à travers l'expérience.

En termes d'**avantages**, on dispose d'une **robustesse** caractéristique des méthodes sub-symboliques, puisque le **réseau** peut s'adapter aux données, tout en bénéficiant de la **précision** et de la **lisibilité** qu'offre la **logique symbolique**. Cela se traduit par un **apprentissage** plus puissant combiné à une **explication** plus accessible. La conclusion ne dépend plus uniquement d'une somme opaque de poids, mais peut être reliée à un **ensemble** comprenant une **règle**, un **concept** ou un **fait**, ce qui en explicite la **signification**. Dans le même paragraphe, on peut souligner que la **cohabitation** de règles symboliques et de flux sub-symboliques exige une **cohérence** de haut niveau. Les règles doivent être appropriées pour décrire les phénomènes observés, et les entités sub-symboliques doivent parvenir à traduire efficacement le **signal** en indices qui **valident** ou **invalident** les **règles**.

Cette **fusion** entre le symbolique et le sub-symbolique au sein du **DSL** instaure un cadre proche de la **cognition synergique**. À l'image de l'esprit humain, le système combine des **concepts** explicites et des **représentations** perceptives implicites, puis les **coordonne** dynamiquement pour aboutir à une interprétation finale.

1.5.7.5. Défis Mathématiques et Implémentation

Dans la poursuite des principes décrits en sections 1.5.7.2 à 1.5.7.4, se pose la question de la **représentation** mathématique des **règles** symboliques à l'intérieur du **Deep Synergy Learning (DSL)** et, plus généralement, les difficultés techniques qu'implique l'introduction de composantes **logiques**. Un premier enjeu consiste à décider comment **encoder** une règle $\mathcal{E}_{\text{rule}}$. Certains travaux préconisent d'associer à chaque règle un **vecteur** ou un **arbre** logique compact, ce qui permet d'attribuer à $\mathcal{E}_{\text{rule}}$ un **nœud** du réseau muni de **paramètres** θ_{rule} . On peut alors définir la **synergie** $S(\mathcal{E}_{\text{rule}}, \mathcal{E}_{\text{feature}})$ de façon à intégrer à la fois la **structure symbolique** (par exemple, un parse-tree) et la **partie numérique** (par exemple, un embedding vectoriel ou un ensemble de features \mathbf{x}_{data}).

Il est possible d'écrire, de manière formelle, une fonction

$$S_{\text{mix}}(\mathcal{E}_{\text{rule}}, \mathcal{E}_{\text{data}}) = F(\text{logic}(\theta_{\text{rule}}), \text{embedding}(\mathbf{x}_{\text{data}})),$$

où **logic** exprime la partie symbolique (règles, connecteurs logiques, etc.), tandis que **embedding** encode la portion sub-symbolique. La fonction F peut évaluer une **co-information** ou un **score** d'adéquation entre la règle et les données. Cette démarche reprend les principes du **DSL** (sections 1.5.7.2 et 1.5.7.3) en donnant un **cadre** unifié pour la synergie entre une entité symbolique et une entité sub-symbolique.

Au fur et à mesure que le **nombre de règles** logiques augmente, on observe que la taille du **réseau** s'accroît, car chaque $\mathcal{E}_{\text{rule}}$ ou $\mathcal{E}_{\text{concept}}$ vient s'ajouter aux nœuds. Les **calculs** de synergie $S_{\text{mix}}(\mathcal{E}_{\text{rule}}, \mathcal{E}_{\text{data}})$ peuvent devenir coûteux si l'on doit évaluer régulièrement de nombreuses règles sur de vastes flux sub-symboliques. Il est alors crucial de recourir à des **techniques** de **parsimonie**, en particulier celles décrites en section 1.4.5, qui permettent de **supprimer** ou de **couper** certaines liaisons $\omega_{i,j}$ dès lors que leur pondération demeure faible. Cette régulation limite la prolifération de liens inutiles, maintenant la **complexité** à un niveau gérable.

Lorsque plusieurs **règles** se **contredisent** au sein du **SCN**, des **tensions** apparaissent. Il est possible que l'une des entités sub-symboliques (par exemple un flux de capteurs) présente une **synergie** positive avec une première règle, mais se heurte à une **incompatibilité** logique avec la seconde. Dans un tel cas, on peut introduire un **score de cohérence** ou une **pénalité** $P(\mathcal{E}_{\text{rule}_1}, \mathcal{E}_{\text{rule}_2})$, reflétant le fait que les règles rule_1 et rule_2 ne peuvent valablement être activées en même temps. On peut alors modifier l'évolution de chaque pondération afin de **réduire** la force de liens conduisant trop souvent à cette contradiction, ou introduire un **terme** correctif dans la mise à jour :

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta [S_{\text{mix}}(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)] - \alpha P(\mathcal{E}_i, \mathcal{E}_j),$$

où α contrôle l'intensité de la pénalité due à la contradiction.

D'un point de vue **implémentation**, les **limites** principales résident dans l'**explosion combinatoire** lorsque l'on souhaite intégrer non seulement des synergies binaires (section 1.4.4) mais également des synergies n-aires (section 1.4.7) et des **connaissances** logiques toujours plus nombreuses. De même, il peut être nécessaire de définir des **règles** plus modulaires ou paramétriques afin d'éviter la création d'une entité symbolique distincte pour chaque énoncé. Les entités pourraient alors coder des **schémas** logiques, associés à des **variables** substituables, pour diminuer le nombre total de nœuds.

Dans l'ensemble, ces **défis** mathématiques et de **mise en œuvre** ne remettent pas en cause le principe fondateur de l'approche hybride (sections 1.5.7.2 et 1.5.7.3), à savoir la possibilité de **fusionner** la rigueur symbolique avec la force d'apprentissage sub-symbolique. Ils incitent plutôt à développer des **algorithmes** de parsimonie adaptés, à concevoir des mécanismes de **cohérence** entre règles, et à prévoir des **heuristiques** ou **méthodes** de clustering pour gérer la croissance du réseau. Cette **fusion** entre le symbolique et le sub-symbolique au sein du **DSL** instaure un cadre proche de la **cognition synergique**. À l'image de l'esprit humain, le système combine des **concepts** explicites et des **représentations** perceptives implicites, puis les **coordonne** dynamiquement pour aboutir à une interprétation finale.

1.5.7.6. Conclusion

L'**incorporation** de composantes **symboliques** ou **cognitives** au sein du **Deep Synergy Learning (DSL)** ouvre un champ d'**IA hybride** dans lequel la **partie sub-symbolique** (traitement de données massives, bruitées et hétérogènes) est **complétée** par une **dimension symbolique** (règles, concepts, ontologies et raisonnement explicite). Cette combinaison se déploie naturellement dans le **Synergistic Connection Network (SCN)** qui, grâce à ses entités adaptatives et à ses liaisons synergiques évolutives, apparaît particulièrement adapté pour réunir et faire coévoluer ces deux approches.

On peut modéliser la **synergie** entre une **règle** R et un **module perceptif** $\mathcal{E}_{\text{data}}$ à l'aide d'une fonction

$$S(\mathcal{E}_{\text{rule}}, \mathcal{E}_{\text{data}}) = f(\theta_{\text{rule}}, \mathbf{x}_{\text{data}}),$$

où θ_{rule} encode l'information symbolique (logique, ontologie...) et \mathbf{x}_{data} représente les **features** sub-symboliques. La pondération $\omega_{\text{rule,data}}(t)$ associée se met alors à jour de façon adaptative, selon

$$\omega_{\text{rule,data}}(t + 1) = \omega_{\text{rule,data}}(t) + \eta [S(\mathcal{E}_{\text{rule}}, \mathcal{E}_{\text{data}}) - \tau \omega_{\text{rule,data}}(t)].$$

Cette dynamique assure que seules les **connexions** réellement bénéfiques (au sens de la synergie mesurée) se consolident, et permet à des **clusters** associant règles symboliques et entités sub-symboliques d'émerger spontanément au sein du réseau.

Ce mécanisme procure de **multiples avantages**. Il **préserve** la capacité d'**apprentissage** sub-symbolique, cruciale pour dompter des données riches et bruitées, tout en **introduisant** des notions symboliques aisément **explicitables** (règles, concepts, axiomes). Cette alliance favorise une **explication** plus claire des décisions, dans la mesure où la conclusion d'un cluster peut être reliée à la **règle** R (logique) et au **signal** \mathbf{x} (apprentissage neuronal). De plus, le DSL conserve ses forces intrinsèques, telles que la **multimodalité**, l'**adaptation** continue face aux variations de distribution et la **résilience** aux données partielles ou contradictoires.

Il faut toutefois noter qu'en **élargissant** la base de **règles** ou de **concepts**, le réseau peut s'agrandir et accroître la **complexité** de calcul, comme évoqué en section 1.5.7.5. Des mécanismes de **parsimonie** ou de **régulation** (limitation des connexions inutiles, seuils dynamiques sur les liaisons) demeurent essentiels pour éviter une explosion combinatoire. On peut également ajouter des **pénalités** ou un **score de cohérence** afin de gérer des contradictions logiques entre plusieurs règles.

En définitive, on obtient ainsi un **DSL** apte à **raisonner** en s'appuyant sur sa **base symbolique**, tout en **apprenant** de manière continue via sa **dimension sub-symbolique**. Cette architecture apparaît comme un **paradigme plus complet**, se rapprochant de la **cognition humaine**, laquelle jongle entre des **concepts** explicites et des **associations** perceptives implicites. Les sections et principes exposés dans ce chapitre (1.5) soulignent la **flexibilité** du DSL et son aptitude à jeter des **ponts** entre la **puissance** d'apprentissage neuronale et la **lisibilité** de la connaissance symbolique.

1.6. Applications Pressenties et Domaines Impactés

Le **Deep Synergy Learning (DSL)**, en tant que paradigme alliant **auto-organisation**, **synergies adaptatives** et **capacité d'évolution continue**, suscite un **intérêt grandissant** dans de nombreux champs de l'intelligence artificielle. Après avoir dressé, dans les sections précédentes, les fondements conceptuels et structurels du DSL (sections 1.1 à 1.5), il est temps de se projeter vers les **applications** et les **domaines** où cette approche pourrait exercer un **impact fort**.

L'architecture **distribuée** et **plastique** du DSL, la possibilité d'**auto-organiser** des entités multimodales ou symboliques, ainsi que sa **tolérance** aux données bruitées, confèrent au DSL un potentiel d'**adaptation** et de **complémentarité** qui dépasse souvent les limites des modèles traditionnels (réseaux neuronaux profonds strictement hiérarchiques ou méthodes symboliques pures). De la **vision artificielle** à la **robotique**, en passant par la **reconnaissance audio**, la **recommandation personnalisée**, la **surveillance** ou encore la **planification industrielle**, le DSL offre des réponses nouvelles, mieux adaptées à la complexité des flux de données et aux exigences d'évolution en temps réel.

Dans cette section (1.6), nous présentons un **éventail** des applications pressenties, non seulement comme une liste d'exemples, mais aussi comme une **cartographie** des domaines que le DSL peut transformer :

- **Vision Artificielle et Reconnaissance d'Objets Complexes (1.6.1)**
- **Analyse Audio et Traitement du Langage Naturel (1.6.2)**
- **Robotique et Systèmes Intelligents Évolutifs (1.6.3)**
- **Recommandation Personnalisée et Systèmes de Décision (1.6.4)**
- **Surveillance, Diagnostic Médical et Anomalies (1.6.5)**
- **Planification et Optimisation dans l'Industrie 4.0 (1.6.6)**
- **Perspectives pour la Recherche Fondamentale en IA Forte (1.6.7)**

Pour chacun de ces thèmes, le **Deep Synergy Learning (DSL)** se révèle prometteur en raison de ses **mécanismes d'auto-organisation**, tels que les **clusters dynamiques**, les **pondérations adaptatives**, l'**intégration de multiples sources** (visuelles, auditives, textuelles, capteurs...) et l'**apprentissage continu** sans dépendance exclusive à la supervision. Nous verrons ainsi comment le DSL peut contribuer à relever certains **défis** actuels de l'IA, tels que la **plasticité**, la **résilience** aux données incertaines, ou encore la **capacité** à modéliser la **complexité** grandissante des problèmes.

1.6.1. Vision Artificielle et Reconnaissance d'Objets Complexes

Parmi les champs qui ont le plus bénéficié de l'essor des réseaux neuronaux profonds, on trouve la **vision artificielle**. Les CNN (Convolutional Neural Networks) ont acquis une renommée considérable pour la détection et la classification d'objets dans des images. Toutefois, les tâches de **reconnaissance** deviennent de plus en plus **subtiles**. Il ne s'agit plus seulement d'identifier un objet, comme un chat ou un chien, mais de **détecter** et **comprendre** des **scènes complexes** dans des environnements changeants, avec des conditions lumineuses variables, des objets partiellement masqués ou des relations multiples entre eux. Le **Deep Synergy Learning** ouvre ici de nouvelles perspectives, dont nous détaillons ci-après les ressorts principaux.

1.6.1.1. Au-delà d'une Hiérarchie Rigide : Clusters d'Entités Visuelles

Dans un CNN classique, les couches de convolution et de pooling s'enchaînent selon un **pipeline** fixe. La première couche détecte des motifs élémentaires comme les bords et les contrastes. La suivante combine ces éléments pour former des structures plus complexes, et ainsi de suite, jusqu'à la classification finale. Bien que très performant, ce mécanisme demeure relativement **rigide**. Il impose un ordre précis selon lequel les "features" visuelles doivent être détectées, sans autoriser la création ou la suppression dynamique de filtres ni de liaisons entre diverses régions de l'image.

Le **Deep Synergy Learning (DSL)** envisage les **features** ou régions d'images comme autant d'**entités** \mathcal{E}_i , chacune portant un **descripteur** lié à un patch, un certain motif ou un vecteur plus abstrait (par exemple issu d'un autoencodeur local). Ces entités visuelles s'**auto-organisent** en **clusters** dès qu'elles identifient une **synergie** notable (proximité de descripteurs, co-occurrences de motifs, etc.). L'**hiérarchie** (ou la "stratification") n'est plus imposée par des couches fixes, mais émerge au contraire de la **coopération** entre entités qui jugent profitable leur rapprochement. On peut par exemple voir se constituer un cluster dédié aux formes circulaires, un autre regroupant les zones de fortes lignes horizontales, et d'autres encore, voire des **macro-clusters** combinant plusieurs de ces groupes si leur fusion devient pertinente.

Cette **dynamique** apporte une plus grande **flexibilité** dans la reconnaissance d'objets complexes ou inattendus. Si un objet n'appartient pas aux catégories standards (par exemple, un objet partiellement masqué ou inconnu), les entités \mathcal{E}_k décrivant ses parties peuvent spontanément **coopérer** avec celles décrivant d'autres objets partiellement similaires. Il se crée alors un **sous-réseau** (cluster) susceptible de fusionner avec un groupement plus vaste, agrandissant le champ d'interprétation.

Pour un modèle plus formel, on peut représenter chaque entité visuelle \mathcal{E}_k par un vecteur $\mathbf{x}_k \in \mathbb{R}^d$. La synergie, qu'elle soit définie via une **distance** (ex. inversée) ou une **co-information**, détermine la valeur $S(\mathcal{E}_k, \mathcal{E}_m)$. Les **pondérations** $\omega_{k,m}$ s'ajustent de façon adaptative au fil des données. Plusieurs entités corrélées finissent par **stabiliser** leurs liaisons internes, formant un **cluster** autour d'un même objet, d'un motif, ou d'une scène particulière. Cette organisation reste **évolutive**. Si le réseau identifie de nouvelles **entités** apportant une information pertinente, les **liaisons** se renforcent, modifiant la structure du **sous-graphe** et intégrant ces découvertes visuelles. La reconnaissance s'opère ainsi de façon plus **organique**, en contrastant avec un pipeline hiérarchique figé. Cette flexibilité permet l'émergence de **regroupements variés** et le **partage de sous-patterns** entre plusieurs objets lorsque cela améliore la synergie du réseau.

1.6.1.2. Multi-Échelle et Robustesse aux Déformations

Dans le traitement d'objets complexes, il est fréquent de rencontrer des **variations** d'échelle (petit/grand dans l'image), des **transformations** géométriques (rotation, vue partielle, ombre) ou une **composition** de sous-parties distinctes. Les réseaux de neurones de type **CNN** gèrent généralement ces phénomènes en pratiquant la **convolution multi-échelle** ou la **data augmentation**, où l'on applique diverses transformations (zoom, rotation...) lors de l'entraînement. Cette stratégie, efficace dans bien des cas, reste toutefois largement paramétrée à l'avance (couches de pooling fixes, patterns choisis pour l'augmentation).

Le **Deep Synergy Learning (DSL)** propose un mécanisme différent. Des **entités** \mathcal{E}_i peuvent décrire un même objet à des échelles ou des angles divers, et leurs **liaisons** $\omega_{i,j}$ se renforcent si la **synergie** (similarité, co-information) confirme qu'il s'agit bien de la **même structure** vue sous différentes perspectives. L'existence d'une entité dédiée à la "version miniaturisée" de l'objet et une autre à la "version grand format" aboutit à un **cluster** commun si elles trouvent un **gain** à leur rapprochement, sans exiger qu'on impose explicitement un niveau de pooling ou un tableau de transformations prédéterminé.

Un premier **exemple** (lié à l'invariance d'échelle) peut être formalisé en introduisant, pour un patch visuel \mathbf{x}_k , une entité $\mathcal{E}_{k,\alpha}$ décrivant ce patch à l'échelle α . Les pondérations $\omega_{(k,\alpha),(k,\beta)}$ s'élèvent si la ressemblance (ou co-information) entre la représentation à l'échelle α et à l'échelle β demeure élevée, ce qui crée un **cluster** multi-échelle au sein duquel ces entités se regroupent.

Un second **exemple** (visant les déformations) consiste à considérer $\mathcal{E}_{k,\theta}$ comme la **version** du patch \mathbf{x}_k après avoir subi une **rotation** θ . Deux angles θ_1 et θ_2 peuvent alors mener à des entités jugées "similaires", renforçant la liaison $\omega_{(k,\theta_1),(k,\theta_2)}$. On peut ainsi constituer un **macro-cluster** stable, reliant divers points de vue d'un même objet, garantissant une forme d'**invariance** à la rotation.

Cette approche confère une **robustesse** aux changements d'échelle et de pose. Le réseau fonctionne sans nécessiter de **couches de pooling** ni de **transformations explicites** ajoutées artificiellement. Le **DSL** permet aux **entités** de **se découvrir** et de **s'associer** spontanément dès qu'un **gain de synergie** est identifié, favorisant ainsi une organisation **adaptative** et **émergente**. L'ajustement se fait localement et **auto-organise** la reconnaissance, en évitant un pipeline rigide, pour composer des structures plus **adaptatives** et résilientes.

1.6.1.3. Interaction avec d'Autres Modalités (Vision + ...)

Dans un cadre multimodal (voir section 1.5.2), la **vision** artificielle s'avère fréquemment complémentée par un **flux audio** ou **textuel**. Le **Deep Synergy Learning (DSL)** n'établit pas de cloison pour la coopération entre l'entité visuelle \mathcal{E}_{vis} et l'entité auditive \mathcal{E}_{aud} . Le fait qu'elles appartiennent à des **modalités distinctes** n'empêche pas leur coopération. Si la **synergie** $S(\mathcal{E}_{\text{vis}}, \mathcal{E}_{\text{aud}})$ est positive, indiquant une **corrélation** ou une **co-information** significative, les **pondérations** $\omega_{\text{vis}, \text{aud}}$ se renforcent progressivement, favorisant la formation d'un **cluster multimodal** adapté à la tâche.

Dans la **reconnaissance** d'objets complexes, un **événement** visuel (comme un objet qui tombe) peut se coupler à un **événement** sonore simultané (bruit d'impact). Si la représentation \mathbf{x}_{vis}

associée à la chute et la représentation \mathbf{x}_{aud} de l'onde sonore se rapprochent au sens de la **coopération**, un **sous-réseau** $\{\mathcal{E}_{\text{patch}}, \mathcal{E}_{\text{featureAudio}}\}$ se constitue. Il se spécialise dans la détection de la scène “objet tombant plus choc acoustique”, ce qui accroît les facultés de **classification** ou de compréhension de la situation (“Cassure”, “Chute de vaisselle”, etc.).

Cette **flexibilité** s'étend à d'autres flux, notamment l'**analyse textuelle**, où des **légendes associées** ou des **mots-clés** issus d'un utilisateur peuvent être pris en compte. Il en résulte la formation de **micro-clusters multi-entités**, où la vision et l'audio coopèrent tout en interagissant avec des entités textuelles. La dynamique d'**auto-organisation** du **DSL** détermine les **groupements** les plus profitables, sans exiger qu'un concepteur définisse un module de fusion dédié pour le couple “vision + audio”. Cela favorise une **intégration** sans couture des diverses modalités, conduisant à une représentation plus riche et plus **résiliente** aux variations d'une modalité prise isolément.

1.6.1.4. Impact sur les Problèmes de Vision Avancés

Dans le cadre de la **vision par ordinateur**, un **Deep Synergy Learning (DSL)** structuré en entités visuelles présente un atout certain pour traiter des scénarios avancés qui outrepassent la simple classification d'images. L'**auto-organisation** autorise l'émergence de **clusters** plus flexibles, allant au-delà du pipeline statique imposé par les méthodes classiques.

Dans la **détection d'objets multiples et partiellement superposés**, par exemple, un réseau traditionnel prévoit souvent un ensemble rigide d'ancrages (bounding boxes) ou un module de segmentation spécifique. Au contraire, le **DSL** peut laisser chaque patch ou région s'**affilier** à des **entités** “chien” et “chaise” même si ces objets se chevauchent dans l'image. Les pondérations $\omega_{i,j}$ augmentent dès que la synergie entre patches associant l'idée de “chien” et l'idée de “chaise” se justifie. Le réseau n'a pas besoin d'un pipeline imposant a priori la détection séparée de chaque catégorie ; la constitution du **cluster** facilite l'identification conjointe de plusieurs objets.

Dans la **segmentation contextuelle**, on peut envisager que certaines entités segmentent l'image en superpixels et d'autres identifient des **contours**. Si leur **synergie** révèle une cohérence — par exemple, les bordures détectées coïncident avec les frontières d'un superpixel —, la liaison ω se renforce. On forme alors un **cluster** correspondant à une “zone cohérente” dans l'image, laquelle pourrait se rattacher à un objet complet ou à une portion d'arrière-plan homogène. Cette approche s'écarte de la segmentation par apprentissage end-to-end, en permettant une coopération plus libre entre entités spécialisées (détection de bordures, regroupement de pixels...) selon la présence d'un **gain**.

Dans un contexte de **vision évolutive en robotique**, un agent dans un environnement changeant (luminosité, décor, nouveaux éléments) doit s'ajuster en continu. Un réseau de neurones traditionnel exigerait soit un réentraînement, soit une réinitialisation partielle. Le **DSL** réévalue localement la **synergie** entre entités visuelles. Les connexions correspondant à d'anciennes conditions, comme une forte luminosité, **diminuent** si elles deviennent obsolètes. En parallèle, de nouvelles **entités** ou **liaisons** émergent pour s'adapter à des variations d'éclairage, des changements d'angles de vue ou l'apparition d'objets inédits. Cette **adaptation** locale maintient la **structure** globale du réseau, sans nécessiter un apprentissage complet de bout en bout, et confère une plasticité plus grande à la **vision** robotique, tout en conservant la possibilité d'exploiter ses acquis antérieurs.

1.6.1.5. Conclusion Partielle : Une Alternative Complémentaire au CNN

Les **CNN** constituent toujours un **vecteur** extrêmement puissant pour la **vision** artificielle, en particulier lorsqu'un large corpus d'exemples annotés est disponible et que l'on vise des tâches de classification ou de détection sur des données stables. Toutefois, la structure d'un **CNN** reste **déterministe**. Chaque couche est **fixe**, et l'apprentissage repose principalement sur une **reconnaissance supervisée**, sans possibilité d'adaptation structurelle dynamique. Le **Deep Synergy Learning (DSL)** apporte une approche plus **organique** et **dynamique**, où les **entités** associées à des *patches* ou des *features* visuelles peuvent librement se **synergiser**, se **séparer**, s'**associer** avec d'autres modalités ou s'adapter aux **conditions** changeantes (nouvelles échelles, bruit, objets inédits).

La **structure** du **Synergistic Connection Network** autorise ainsi une **auto-organisation** continue, évitant la rigidité d'un pipeline imposé pour la vision. On peut introduire de nouvelles entités $\mathcal{E}_{\text{nouvelle}}$ si un objet inconnu apparaît ou si l'on veut traiter un autre type de patch visuel, ce qui met à jour localement les connexions ω . Le **DSL** peut s'**hybrider** avec un **CNN** en exploitant les **descripteurs** extraits d'une couche intermédiaire comme entités \mathcal{E}_{CNN} . La **synergie** entre ces entités reflète alors leur **complémentarité**, permettant une structuration plus dynamique et adaptative des informations issues du réseau convolutionnel.

Pour la **reconnaissance d'objets complexes** dans des scènes variées (éclairage fluctuant, angles de vue inédits, transformations géométriques), le **DSL** fournit un **cadre** mathématique structuré, avec des **équations** décrivant l'évolution des liaisons $\omega_{i,j}$ et des mesures de synergie multi-échelle (ou non linéaire). Il enrichit la **vision** d'une dose de **plasticité** et de **robustesse**, sans empêcher l'usage de CNN comme base de features. Les **clusters** qui émergent dans ce réseau renvoient à des regroupements de patches ou de motifs liés par leur **gain mutuel**, offrant un point de vue plus distribué et collaboratif qu'un pipeline hiérarchique unique. Au final, le **DSL** s'érige en complément précieux aux **CNN** pour atteindre une **vision** flexible, résiliente et progressivement évolutive.

1.6.2. Analyse Audio et Traitement du Langage Naturel

Après avoir vu comment le **Deep Synergy Learning (DSL)** peut apporter une flexibilité et une auto-organisation en **vision artificielle** (1.6.1), il est naturel de s'intéresser aux **données audio** et au **langage naturel**, deux domaines où la richesse des signaux et la variabilité contextuelle sont particulièrement élevées. Dans un réseau neuronal profond classique, les approches d'**analyse audio** (speech recognition, classification de sons, détection d'événements acoustiques) et de **traitement du langage naturel** (NLP) (analyse sémantique, traduction, question-réponse) demeurent puissantes, mais reposent souvent sur des modèles séparés (RNN, LSTM, Transformers) et des schémas d'apprentissage supervisé imposant de gros volumes de données annotées.

Le **DSL**, en revanche, repose sur une **auto-organisation** où chaque entité \mathcal{E}_i (qu'elle soit issue de données audio ou textuelles) s'insère dans un **réseau synergique** plus large. Cette approche permet une **fusion plus fluide** des informations acoustiques et linguistiques, facilitant une **adaptation**

continue aux variations telles que le bruit, les accents ou les registres de langage. De plus, elle favorise une **cohabitation** entre les représentations symboliques et sub-symboliques, rendant possible l'intégration de règles logiques dans des tâches linguistiques plus complexes.

Cette sous-section (1.6.2) détaille en quoi le DSL modifie l'approche habituelle de l'analyse audio et du traitement du langage, et en quoi il répond à des défis d'**hétérogénéité**, de **bruit**, de **variabilité linguistique** et de **scénarios évolutifs**.

1.6.2.1. Analyse Audio : Entités Sonores et Synergie Adaptative

Dans un **réseau** neuronal **classique** pour l'audio (CNN ou RNN appliqué aux spectrogrammes), on spécifie dès le départ un schéma d'extraction de *features* (MFCC, log-mel, etc.), puis l'on optimise un classifieur (par exemple, pour la reconnaissance de phonèmes ou la classification de sons). Le **Deep Synergy Learning (DSL)** introduit une logique plus **organique** en laissant des **entités** $\mathcal{E}_{\text{aud},k}$ se constituer pour chaque **fenêtre** ou **segment** temporel (frames audio), pour un **spectre** fréquentiel spécifique, ou pour un **embedding** plus abstrait (mots, bruits, etc.). Ces entités se regroupent en **clusters** dès lors qu'elles observent un **gain** à coopérer, qu'il s'agisse de la ressemblance spectrale, de la proximité temporelle ou de la récurrence d'un même motif acoustique.

Si on modélise deux segments audios $\mathbf{a}_i, \mathbf{a}_j \in \mathbb{R}^d$, leur synergie peut être définie par une **distance inversée** ou une **similitude**, telle que

$$S(\mathcal{E}_i, \mathcal{E}_j) = \exp(-\gamma \|\mathbf{a}_i - \mathbf{a}_j\|^2),$$

pour un certain $\gamma > 0$. La **pondération** $\omega_{i,j}(t)$ qui relie les entités \mathcal{E}_i et \mathcal{E}_j se met alors à jour selon la règle adaptative (voir section 1.4.5). Lorsque deux **fenêtres** ou segments audio affichent une forte similarité (timbre commun, bruit de fond identique, même locuteur), leurs liaisons $\omega_{i,j}$ se **renforcent**, initiant ainsi la création d'un **cluster** rassemblant ces entités, et permettant de **repérer** un même motif acoustique. Cette **auto-organisation** se substitue aux approches classiques qui segmentent l'audio selon des règles prédéfinies. Plutôt que d'imposer une structure rigide, le réseau ajuste dynamiquement ses **liaisons** en fonction de la synergie perçue, permettant une **segmentation** ou une **agrégation** adaptative des informations.

De plus, si un **environnement** se modifie (bruit important, micro défaillant, accent inhabituel), des entités précédemment utiles peuvent perdre leur **synergie** avec les autres composantes, faisant décroître $\omega_{i,j}$. L'entité se retrouve alors **isolée** et n'influence plus le cœur du réseau. Inversement, lorsqu'une nouvelle configuration sonore émerge (baisse du bruit, nouveaux sons caractéristiques), d'autres **connexions** ω apparaissent ou se renforcent, configurant de nouveaux **clusters** sonores. Ainsi, la **reconnaissance** de motifs acoustiques (sirènes, voix enfantines, grondements) se met en place de manière **continue**, sans qu'on doive réentraîner de bout en bout un réseau rigide. Le **DSL** s'adapte localement, confortant l'idée d'un système plus **résilient** et **évolutif** dans le traitement de l'audio.

1.6.2.2. Traitement du Langage Naturel : Entités Lexicales et Sémantiques

Dans les méthodes **classiques** de **traitement du langage naturel** (NLP), des modèles (RNN, LSTM, Transformers) s'appuient sur des **embeddings** (word2vec, GloVe, BERT, etc.) et des mécanismes d'attention ou de convolution. Bien qu'ils aient atteint d'excellents résultats (traduction, question-réponse, résumé, etc.), ils s'avèrent fortement dépendants d'un **entraînement** supervisé ou auto-supervisé de grande ampleur, et peinent à **intégrer** de la **logique** ou du **symbolique** sans recourir à des mécanismes additionnels. De plus, la **structure** de ces modèles demeure souvent linéaire (ou en arbre pour l'analyse syntaxique), imposant le traitement séquentiel des tokens.

Le **Deep Synergy Learning (DSL)** propose une alternative plus **flexible**. Au lieu d'imposer une séquence rigide de tokens à travers un pipeline (embedding + attention + couches finales), on peut **déployer** un **réseau** d'entités couvrant les différentes dimensions du texte. Chaque entité \mathcal{E}_i peut représenter un **mot** (ou token) muni d'un embedding \mathbf{w}_i , un **nœud syntaxique** extrait d'un parse tree (rôle de sujet, verbe, complément), ou un **concept sémantique** (topic, entité nommée, etc.). Ces entités s'**auto-organisent** si leur synergie $S(\mathcal{E}_i, \mathcal{E}_j)$ s'avère suffisamment élevée pour justifier un **regroupement**.

On peut alors voir naître des **clusters** lexical-sémantique regroupant des mots liés à un même champ lexical, un **verbe** et ses **compléments**, ou plusieurs **concepts** partagés par diverses phrases. Plutôt qu'une **arborescence figée**, comme dans une analyse syntaxique classique, le **SCN** adopte une **structure fluide**. Lorsqu'un mot présente plusieurs sens possibles, la **synergie** avec d'autres entités textuelles ou contextuelles oriente naturellement son interprétation. Ce mécanisme permet une **désambiguïsation dynamique**, où le sens retenu dépend des relations de synergie établies en temps réel au sein du réseau. De même, des liaisons $\omega_{i,j}$ peuvent se **rompre** si un ensemble de mots cesse d'être pertinent dans un nouveau contexte textuel. L'**auto-organisation** rend donc la structure du **réseau** ajustable, permettant d'incorporer ou de dissoudre des sous-groupes de tokens ou de concepts selon qu'un **gain** mutuel se manifeste, ouvrant ainsi la voie à une **compréhension** plus modulable du langage.

1.6.2.3. Couplage Audio-Texte et Auto-Fusion Synergiques

De nombreux systèmes doivent traiter simultanément un **flux audio** (voix, intonations, bruits) et un **texte** associé (transcriptions, sous-titres). Le **Deep Synergy Learning (DSL)** s'avère particulièrement bien adapté à ce scénario, car il ne fixe pas a priori la manière de "joindre" l'audio et le texte. Au lieu d'un alignement imposé par un algorithme dédié, on laisse le **Synergistic Connection Network (SCN)** réguler les liaisons ω entre des **entités** \mathcal{E}_{aud} et des **entités** \mathcal{E}_{txt} . Si la **co-information** ou une autre mesure de **similarité** suggère qu'un certain segment sonore correspond à un certain segment textuel, la pondération $\omega_{\text{aud}, \text{txt}}$ se renforce, conduisant à la création d'un **cluster** multimodal.

Parmi les entités audio, on peut distinguer des segments temporels ou des frames associées à une **fenêtre** spectrale spécifique. Du côté textuel, on dispose d'entités représentant des **mots**, des **tokens**, voire des syntagmes plus longs. La mise à jour des liens $\omega_{\text{aud}, \text{txt}}$ procède alors par la règle adaptative (section 1.4.5), détectant les **paires** qui coïncident fréquemment. Ainsi, lorsqu'un

segment audio s'aligne régulièrement sur une suite de mots, on observe un gain synergique. On forme donc un **cluster** englobant les parties orales et leurs correspondances textuelles, sans qu'un schéma d'alignement fixe soit nécessaire.

Un tel mécanisme confère une **robustesse** notable. Si des segments audio se montrent inintelligibles (forte perturbation, accent trop marqué), la **synergie** s'affaiblit, et les liaisons avec les entités textuelles cessent de croître, isolant ces segments pour ne pas polluer l'ensemble. Le réseau **accepte** ainsi l'hétérogénéité de la qualité sonore et s'auto-adapte. En outre, cette logique **permet de découvrir** des co-occurrences ou des associations inhabituelles, par exemple un mot (ou un motif lexical) qui se répète chaque fois qu'un type particulier de sonorité apparaît. Le **DSL** parvient donc à fusionner audio et texte de manière plus **flexible** que ne le ferait un pipeline d'alignement rigide, et il favorise l'intégration de nouvelles données ou la détection de nouveaux mots/sons sans devoir procéder à un réentraînement ou une refonte exhaustive du modèle.

1.6.2.4. Évolutions Possibles et Défis Mathématiques

Dans l'application du **Deep Synergy Learning (DSL)** à l'audio, au texte ou à des mélanges de modalités, plusieurs difficultés et pistes d'extension apparaissent.

Un premier enjeu concerne la **complexité**. Comme en vision (section 1.6.1.4), on peut générer un grand nombre d'entités audio (fenêtres temporelles, frames) ou textuelles (tokens, syntagmes). Le **graph** qui en résulte peut ainsi atteindre une taille considérable, rendant la mise à jour des pondérations $\omega_{i,j}$ et l'estimation des synergies coûteuses. Il est donc souvent nécessaire de recourir à des **stratégies** parcimonieuses pour limiter la prolifération de liens, comme échantillonner périodiquement les entités, fusionner les entités jugées redondantes ou fixer un seuil ω_{\min} qui supprime les connexions trop faibles (section 1.4.5).

En second lieu, on notera que la **modélisation n-aire** peut se révéler essentielle. Les **synergies** binaires (audio–audio, texte–texte ou audio–texte) ne suffisent pas toujours à capturer des phénomènes impliquant trois (voire davantage) entités. Il se peut, par exemple, que l'union de deux fragments audio et d'un groupe de mots textuels produise une information inatteignable avec des paires isolées. La **synergie n-aire** (section 1.4.7) permet de refléter ces complémentarités plus complexes, au prix d'un accroissement de la difficulté mathématique, car il faut évaluer l'information ou la similarité simultanément sur plusieurs variables.

Enfin, la possibilité d'**hybridation symbolique** apparaît particulièrement intéressante dans le domaine du **NLP** avancé (raisonnement, question-réponse complexes). Comme évoqué en section 1.5.7, on peut introduire des **règles logiques** ou des **entités** symboliques, lesquelles interagissent avec les **entités textuelles**. La synergie se définit alors par une fonction reliant le sens des mots à des conditions logiques ou des assertions formelles. Si cette intégration accroît encore la complexité, elle apporte un surcroît de **capacité cognitive**, autorisant un raisonnement plus poussé et assurant une **explicabilité** renforcée, dans l'esprit d'une IA **neuro-symbolique** où la dynamique des pondérations répond à la fois aux critères sub-symboliques (similarité, co-information) et aux axiomes symboliques.

1.6.2.5. Conclusion Partielle : un Cadre Évolutif pour l'Audio et le Langage

Le **Deep Synergy Learning (DSL)** s'affirme comme une **approche** ou un **complément** aux architectures traditionnelles (CNN, RNN, Transformers) dans l'analyse de **flux audio** et le **traitement du langage naturel**. Contrairement aux pipelines classiques, le **DSL** ne se fonde pas sur un découpage préalable ou un alignement imposé, mais laisse les entités acoustiques ou lexicales s'**auto-organiser** dès que leur synergie l'indique. Cette approche apporte plusieurs **bénéfices** en matière de **segmentation adaptative** et de **fusion multimodale**.

L'**auto-organisation** ajuste dynamiquement la segmentation et l'alignement, consolidant les pondérations entre segments audio et unités textuelles dès qu'une corrélation ou une complémentarité apparaît. Contrairement aux méthodes traditionnelles qui imposent un découpage temporel rigide, le **réseau** ajuste ses connexions de façon fluide, ce qui le rend **robuste** face aux variations acoustiques (bruit, accents, altérations de source) et linguistiques (polysémie, registres de langage). Lorsqu'une information devient incertaine ou bruitée, les liaisons s'atténuent d'elles-mêmes, évitant un réapprentissage global.

La **fusion** entre l'audio et le texte se fait de manière **naturelle**, sans exiger un alignement strict entre un segment oral et sa transcription. Le **DSL** découvre spontanément les correspondances pertinentes en évaluant la **synergie** entre unités sonores et textuelles, indépendamment d'un cadre prédéfini. Cette flexibilité permet d'intégrer **de nouveaux sons, du vocabulaire inédit ou des règles symboliques** sans nécessiter une reconstruction complète du système. Il suffit d'introduire de nouvelles **entités** dans le graphe, et la mise à jour des pondérations s'effectue automatiquement.

Ces caractéristiques favorisent plusieurs **applications évolutives** en traitement du langage et de l'audio. La **reconnaissance vocale** en milieu bruité devient plus robuste grâce à l'**auto-organisation**, qui ajuste les pondérations en fonction de la qualité du signal. La **compréhension** des conversations multimodales, intégrant audio et vidéo, se fait sans imposer un **pipeline** d'alignement rigide. La **traduction** et la **détection** d'anomalies audio-linguistiques bénéficient également de cette structure **distribuée et dynamique**, qui ajuste les liens entre entités selon la **synergie** mesurée.

Ce cadre repose sur des **règles d'auto-organisation** qui adaptent les pondérations et façonnent des **clusters**, tout en s'appuyant sur des critères de **synergie** tels que la distance et la co-information. L'ensemble favorise un fonctionnement plus **flexible**, moins dépendant d'un entraînement supervisé massif, et davantage axé sur une **évolution continue** où les entités coopèrent localement pour optimiser l'interprétation des données.

1.6.3. Robotique et Systèmes Intelligents Évolutifs

Les concepts d'**auto-organisation**, de **synergie** et de **coopération** entre entités, propres au **Deep Synergy Learning (DSL)**, prennent tout leur sens dans le domaine de la **robotique** et des **systèmes intelligents**. En effet, la robotique moderne repose sur l'**intégration** de capteurs variés tels que la vision, l'audio, la pression ou la position, chacun fournissant des données essentielles. Les **actions** doivent être coordonnées avec précision, qu'il s'agisse de mouvements, de manipulations ou d'interactions, tout en tenant compte des changements de l'environnement. L'**adaptation continue** est primordiale pour ajuster la dynamique du robot face aux imprévus, aux variations de l'espace

ou aux perturbations extérieures. Enfin, la **planification** et la **prise de décision en temps réel** exigent une gestion efficace de l'**incertitude** et du **bruit**, rendant indispensable une approche flexible et évolutive.

Cette sous-section (1.6.3) montre comment le **DSL**, grâce à ses mécanismes de **pondérations adaptatives** et à ses **clusters auto-organisés**, peut être particulièrement utile à la **robotique** et aux **systèmes intelligents**. Elle décrit :

- L'**approche distribuée** que le DSL propose pour les robots multi-capteurs,
- La **coévolution** des entités sensorielles et motrices,
- La capacité à **s'auto-réorganiser** lorsqu'un robot ou un système intelligent découvre de nouveaux modules (nouveaux effecteurs, nouveaux contextes, etc.),
- Des exemples d'applications, depuis les **robots collaborateurs** jusqu'aux **systèmes autonomes** en changement permanent.

1.6.3.1. Robotique Multi-Capteurs : une Structure Synergique

Dans un contexte robotique, un système typique réunit plusieurs **capteurs** disposés sur la plateforme, pouvant inclure une **caméra** (analyse visuelle), un **LIDAR** ou un **radar** (cartographie de l'environnement), des **capteurs de pression** ou de **toucher**, des **capteurs inertiels** (IMU), un **microphone** pour la partie audio, etc. Les approches traditionnelles (notamment basées sur ROS et des architectures préconçues) tendent à définir un **pipeline** de fusion, auquel chaque capteur envoie ses données, puis une **couche** de décision. Cette rigidité suppose un schéma de liaison explicite entre capteurs et modules supérieurs, imposant par exemple un module de traitement conjoint image-LIDAR si l'on veut superposer un nuage de points laser et un flux de caméra.

Le **Deep Synergy Learning (DSL)**, quant à lui, permet d'instituer un **ensemble** d'entités $\{\mathcal{E}_{\text{cam}}, \mathcal{E}_{\text{lidar}}, \dots\}$ correspondant chacune à une source sensorielle ou à un bloc fonctionnel plus abstrait. Chaque **entité** du système évalue la **synergie** qu'elle entretient avec les autres. Par exemple, si la caméra \mathcal{E}_{cam} et le LIDAR $\mathcal{E}_{\text{lidar}}$ détectent fréquemment le même obstacle, leur liaison se **renforce** progressivement, ce qui entraîne une consolidation de la pondération $\omega_{\text{cam}, \text{lidar}}$. Cette dynamique adaptative permet aux capteurs de **s'ajuster** en fonction de leur cohérence mutuelle, optimisant ainsi la perception et la prise de décision du système. Inversement, lorsqu'un capteur inertiel s'avère temporairement *bruité* (information incohérente par rapport aux autres lectures), ses connexions décroissent naturellement jusqu'à s'affaiblir et, le cas échéant, s'isoler.

Grâce à ce principe d'**auto-organisation**, il devient possible qu'un **cluster** multimodal émerge pour caractériser une situation telle qu'un "terrain accidenté" perçu simultanément par la caméra (textures complexes), le LIDAR (irrégularités de distance) et peut-être un micro détectant des bruits anormaux. Ce cluster se **dissout** ensuite lorsque le robot quitte la zone problématique ou que la configuration change. Le **réseau** d'entités agit donc de manière plus **dynamique** et **distribuée**, sans imposer un module de fusion central dédié à chaque couple de capteurs. Les liaisons $\omega_{i,j}$ se forment ou se dissolvent localement, en fonction de la **valeur ajoutée** mesurée à travers la synergie

entre entités, conférant au système une **résilience** plus élevée, qu’il s’agisse de gérer un capteur défaillant, l’arrivée d’un nouveau module, ou un changement inattendu dans l’environnement.

1.6.3.2. Action, Effecteurs et Synergie Motrice

En robotique ou dans tout système intelligent disposant de **modules d’action** (moteurs, articulations, roues, pinces), on adopte généralement un paradigme où un **contrôleur** (PID, MDP, RL, etc.) reçoit un **état** en provenance des capteurs et génère des **commandes**. Le **Deep Synergy Learning (DSL)** aborde la question autrement, en considérant les effecteurs comme des **entités** $\mathcal{E}_{\text{motrice}}$ au même titre que les capteurs. Par exemple, chacune des articulations d’un bras, ou la rotation d’une roue, ou la position d’une pince peut être décrite par un vecteur $\mathbf{x}_{\text{motrice}}$. La **coopération** avec les entités sensorielles s’établit dès lors que la **synergie** $S(\mathcal{E}_{\text{capteur}}, \mathcal{E}_{\text{motrice}})$ s’avère positive, incitant à **renforcer** la pondération $\omega_{\text{capteur}, \text{motrice}}$.

Un bras articulé qui coordonne son mouvement avec un flux visuel voit, par exemple, la pondération $\omega_{\text{cam}, \text{bras}}$ croître si la caméra confirme que la pince se trouve dans la bonne position. À l’inverse, un joint dysfonctionnel ou mal calibré cesse d’apporter un réel gain, ce qui fait chuter la **synergie**, rendant son influence moindre dans la stratégie globale. Le **SCN** assure ainsi une forme d’**adaptation** où les effecteurs pertinents coopèrent étroitement avec les capteurs produisant des informations utiles, sans qu’il soit nécessaire de configurer explicitement un contrôleur central devant orchestrer toutes les modalités.

On peut également insérer un **signal** de récompense ou un indicateur de performance R qui module la synergie. Si la coopération entre un capteur $\mathcal{E}_{\text{capteurA}}$ et un effecteur $\mathcal{E}_{\text{motriceB}}$ accroît ce score (par exemple, la manipulation d’un objet est mieux réalisée, ou la trajectoire est plus stable), la **pondération** $\omega_{\text{capteurA}, \text{motriceB}}$ s’en trouve augmentée. Cette logique correspond à un **processus** de “learning by synergy”, sans avoir à recourir à une formalisation en MDP (Markov Decision Process) rigide. Les entités effectrices et sensorielles **s’auto-organisent** localement via la mise à jour des connexions ω , découvrant progressivement quelles combinaisons capteur–action se révèlent fructueuses pour atteindre l’objectif.

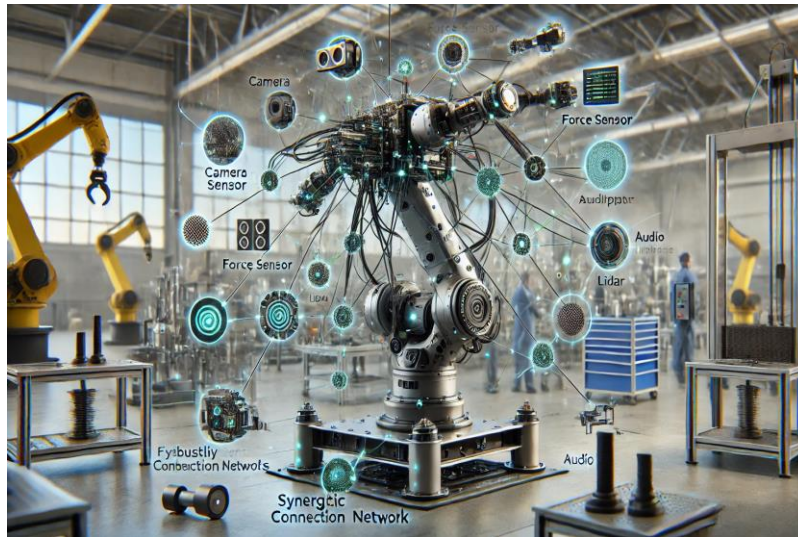
1.6.3.3. Adaptation Continue et Évolution Modulaire

Cette **flexibilité** structurelle constitue un atout majeur du **Deep Synergy Learning (DSL)**. Lorsqu’un robot reçoit un **nouveau capteur** ou un **nouvel effecteur**, ou qu’un module obsolète est retiré, il suffit d’**ajouter** ou de **supprimer** l’entité correspondante $\mathcal{E}_{\text{nouveau}}$ au sein du **Synergistic Connection Network (SCN)**. La **dynamique adaptative du réseau ajuste automatiquement les liaisons synergiques**. Les **nouvelles connexions se renforcent si elles apportent un gain, tandis que les anciennes, devenues inutiles, s’atténuent sans nécessiter de réapprentissage exhaustif**. Les **pondérations** ω_{nouveau} suivent la **dynamique habituelle (voir section 1.4.5)** et se renforcent ou se **dissolvent** en fonction de la **synergie** détectée. Il n’est donc pas requis de recompiler l’ensemble du pipeline ou de procéder à un **réapprentissage** complet d’un grand réseau neuronal. La **logique** adaptative se borne à évaluer localement la valeur ajoutée de la nouvelle entité, puis à l’intégrer dans un **cluster** pertinent si besoin.

Cette **propriété** ne se limite pas à la **robotique** physique. Dans des **systèmes logiciels** complexes, par exemple dans l'administration de centres de données ou d'architectures distribuées, le principe reste valable. Chaque microservice, base de données, module de monitoring, etc., peut être considéré comme une **entité** $\mathcal{E}_{\text{service}}$. L'arrivée d'un nouveau service, la suppression (ou la panne) d'un service existant, se modélise par l'**ajout** ou le **retrait** d'entités correspondantes, et les **liaisons** ω se régulent selon la **coopération** réelle (échange d'informations, dépendances). Les **clusters** qui en résultent rassemblent les **services** ayant des affinités ou co-occurrences fortes, autorisant une **répartition** et une **coordination** modulaires.

Ce **principe d'auto-organisation** évite la centralisation rigide d'un pipeline complet ou la révision constante d'un réseau monolithique. Le **SCN** s'ajuste **spontanément**. Les entités fiables sont **préservées**, tandis que celles devenues obsolètes voient leurs liaisons $\omega_{i,j}$ s'atténuer. Lorsqu'une **nouvelle entité** apparaît, le réseau **restructure** ses connexions pour maximiser la **synergie**, sans nécessiter un **réapprentissage complet**. Cette démarche s'inscrit dans l'esprit du **lifelong learning**, où le **réseau** lui-même façonne en continu sa **topologie** et maintient un équilibre entre la conservation de connaissances antérieures et l'ajout de briques inédites.

1.6.3.4. Illustration : Robot Collaboratif Multicapteurs



Un **robot collaboratif** (cobot) dans un environnement industriel peut disposer de multiples capteurs et actionneurs. Par exemple, on considère une **caméra** pour localiser la pièce à assembler, un **capteur de force** dans la pince pour jauger la préhension, un **LIDAR** pour détecter la distance aux opérateurs humains et éviter les collisions, un **module audio** pour recevoir des instructions vocales, ainsi que les actionneurs comprenant la **pince**, le **bras articulé** et un **moteur** de déplacement sur plateforme.

Dans un **Synergistic Connection Network (SCN)** selon le **Deep Synergy Learning (DSL)**, chaque composant est représenté comme une **entité** : \mathcal{E}_{cam} , $\mathcal{E}_{\text{force}}$, $\mathcal{E}_{\text{lidar}}$, $\mathcal{E}_{\text{audio}}$ pour les capteurs, et $\mathcal{E}_{\text{pince}}$, $\mathcal{E}_{\text{bras}}$, $\mathcal{E}_{\text{moteur}}$ pour les effecteurs. Chacune évalue localement sa **synergie** avec les autres, de sorte que les liens $\omega_{i,j}$ se **renforcent** s'il apparaît un réel bénéfice à leur coopération. Lorsque la caméra coopère régulièrement avec la pince — par exemple, parce que la zone repérée par

l'image correspond effectivement à la position où la pince intervient —, la liaison $\omega_{\text{cam,pince}}$ se consolide, engendrant un **cluster** local regroupant ces entités {cam, pince}. Si, au contraire, l'audio se révèle trop bruyant, son synergie avec le reste du système diminue et ω_{audio} tombe en deçà du seuil de pertinence ; l'entité audio se retrouve alors **isolée** ou peu contributive jusqu'à ce que la qualité du signal s'améliore.

On peut intégrer un **signal de récompense** afin de consolider les combinaisons sensorielles et motrices les plus efficaces (par exemple, si le cobot réussit à saisir et assembler la pièce sans heurter de zones sensibles, on valorise cette configuration). Les entités impliquées voient leurs pondérations internes augmenter, ce qui renforce la probabilité de réutiliser ce mode de coopération. L'ensemble constitue un **réseau** qui s'**auto-adapte**, insérant ou affaiblissant des liens selon l'utilité détectée pour la tâche courante. Cette logique évite de reconstruire un pipeline complet ou de réentraîner un vaste réseau en cas de modifications (ajout d'un nouveau capteur, panne de l'actuel, variation des conditions d'assemblage). Les **clusters** multimodaux (vision + pince + capteur de force, par exemple) forment des **sous-réseaux** cohérents, naturellement conçus pour la tâche à effectuer.

1.6.3.5. Défis et Perspectives de Recherche

Lorsqu'on applique le **Deep Synergy Learning (DSL)** à la robotique multi-capteurs et multi-effecteurs, plusieurs difficultés et voies d'extension émergent, tant sur le plan pratique que théorique.

Un premier défi concerne la **complexité combinatoire**. Dans un robot doté de multiples capteurs (caméra, LIDAR, IMU, microphone...) et d'autant de modules d'action (divers joints, roues, pinces), le **graphe** des pondérations ω peut rapidement croître en taille, rendant la mise à jour de chaque lien difficile à maintenir en **temps réel**. Des stratégies de **parsimonie** (seuils éliminant les liaisons trop faibles) ou des mises à jour strictement **locales** (limiter le calcul aux entités concernées par une nouvelle donnée) se révèlent indispensables pour garantir la **scalabilité**.

Un deuxième point aborde les **systèmes multi-robots**. Lorsqu'une flotte (drones, véhicules, robots industriels) doit coopérer, on peut concevoir un **réseau synergique** à plus grande échelle, où chaque robot forme un **sous-graphe** d'entités sensorielles et motrices. Les **synergies** inter-robots introduisent alors des liaisons ω entre entités de différents appareils. Il devient possible d'imaginer une **co-organisation** sur l'ensemble de la flotte, favorisant la dynamique d'émergence de clusters inter-robots dès qu'un gain mutuel apparaît (partage de capteurs, répartition de tâches). L'envergure de ce problème soulève d'importants défis de **communication** et de gestion **distribuée** des liens, car on ne peut pas nécessairement centraliser toutes les informations.

La **sécurité** et la **robustesse** forment également un axe de recherche majeur. En robotique, il est impératif de respecter des contraintes physiques (éviter la collision, maintenir une distance de sécurité). Dans le **DSL**, on peut introduire des pénalités ou des **contraintes** pour réduire la pondération $\omega_{i,j}$ lorsqu'une configuration s'avère dangereuse ou enfreint des invariants mécaniques ou de sûreté. De cette manière, l'**auto-organisation** se soumet à des garde-fous, garantissant un comportement sûr même lorsque la structure du réseau évolue.

Enfin, l'idée d'**hybridation symbolique** (section 1.5.7) se transpose à la robotique avancée et aux tâches haut niveau. Les entités symboliques, représentant des **règles** ou des **plans**, interagissent avec les entités sensorielles et motrices. La synergie entre la **logique** (plan d'action) et la **perception** (reconnaissance visuelle, retour de force) réorganise le graphe de manière dynamique. Cette interaction aboutit à une planification distribuée et **cognitive**, où la coopération se produit entre le symbolique (ordre de mission) et le sub-symbolique (capteurs, effecteurs). L'ensemble de ces pistes ouvre un **vaste** espace de recherche pour faire évoluer la robotique vers des systèmes de décision auto-organisés, hautement modulaires et résilients.

1.6.4. *Recommandation Personnalisée et Systèmes de Décision*

Au-delà de la robotique (1.6.3), le **Deep Synergy Learning (DSL)** trouve également une place de choix dans les **systèmes de recommandation** et plus largement les **systèmes de décision**. Dans ces domaines, il s'agit souvent de traiter des **profils utilisateurs**, des **contenus** (produits, articles, films, musiques) ou des **options** (plans d'action, configurations possibles), pour aboutir à des **conseils** ou des **choix** pertinents. Les approches traditionnelles, comme le **collaborative filtering**, les **réseaux neuronaux de recommandation** ou les **arbres de décision**, reposent sur des structures fixes. Un modèle est défini, puis entraîné de manière plus ou moins supervisée pour calculer des **scores** de **préférence**. Le **DSL**, par sa **dynamique d'auto-organisation** et sa **prise en charge** des liens synergiques, propose une voie plus **flexible** et **adaptive**. Il peut gérer à la fois :

- Des **entités “utilisateurs”** $\{\mathcal{E}_u\}$ représentant différentes personnes (ou profils),
- Des **entités “contenus”** $\{\mathcal{E}_c\}$ décrivant articles, produits, médias, etc.,
- Des **entités “contextes”** (saisons, heures, tendances récentes),
- Éventuellement des **entités symboliques** (règles de business logic, préférences explicites) si on veut un système cognitif plus avancé.

Nous allons voir en détail comment ces entités interagissent et comment le DSL peut aider à **adapter** les recommandations au fil du temps, en favorisant les **clusters** (ou micro-réseaux) les plus cohérents, et en révisant les connexions obsolètes.

1.6.4.1. *Entités Utilisateurs et Contenus : un Graphe Synergique*

Les **systèmes de recommandation** s'appuient classiquement sur une **matrice M** dont les lignes représentent les utilisateurs $\{u \in U\}$ et les colonnes les contenus $\{c \in C\}$. Chaque entrée $M_{u,c}$ correspond à l'intérêt ou la note attribuée par l'utilisateur u au contenu c . Les approches courantes (collaborative filtering, factorisation matricielle, autoencodeurs de complétion...) consistent alors à factoriser **M** ou à prédire les valeurs manquantes par un entraînement supervisé ou semi-supervisé. Cette vision reste **matricielle** et se limite généralement au couple **utilisateurs-contenus**, sans intégrer d'autres **entités** susceptibles d'influencer la recommandation.

Dans le **Deep Synergy Learning (DSL)**, on substitue ou complète cette approche en définissant un **graphe** évolutif. On introduit d’abord un ensemble d’**entités** $\{\mathcal{E}_u\}_{u \in U}$ représentant les utilisateurs et un autre $\{\mathcal{E}_c\}_{c \in C}$ pour les contenus (articles, films, produits, etc.). Si nécessaire, on y ajoute aussi des **entités contextuelles** $\{\mathcal{E}_{\text{time}}, \mathcal{E}_{\text{location}}, \dots\}$ lorsqu’on souhaite prendre en considération la dimension temporelle ou spatiale. Les **liaisons** $\omega_{u,c}(t)$ entre un utilisateur \mathcal{E}_u et un contenu \mathcal{E}_c évoluent selon une **règle adaptative** (voir section 1.4.5). Elles se renforcent lorsque la **synergie** $S(\mathcal{E}_u, \mathcal{E}_c)$ est élevée et déclinent en l’absence d’utilité mutuelle. La fonction S repose sur des critères tels que la **similarité** d’intérêts, la **co-information** d’usage ou un gain mesurable, comme le nombre de clics ou de vues.

Grâce à ce schéma, on assiste à la formation spontanée de **clusters** regroupant simultanément des utilisateurs et des contenus. Un **cluster** peut ainsi rassembler plusieurs amateurs de rock, divers albums rock et, s’il existe des entités de contexte, certaines soirées “concert” associées, révélant une forte **co-occurrence** ou un alignement d’intérêts. Comparé à la matrice **M** — qui se borne aux couples (u, c) — le **SCN** tolère l’ajout de toute entité auxiliaire. Il devient possible de **fusionner** des groupes si une synergie apparaît entre leurs membres, ou d’**isoler** un contenu dont l’intérêt décroît soudain auprès des utilisateurs concernés. En outre, l’arrivée d’un nouvel utilisateur $\mathcal{E}_{\text{nouveau}}$ ou d’un nouveau contenu $\mathcal{E}_{\text{item}}$ se gère en insérant simplement l’entité, puis en laissant la synergie ajuster les liens $\omega_{\text{nouveau}, \cdot}$. Le réseau se **restructure** localement, sans imposer une refonte globale comme le ferait une ré-analyse complète de la matrice **M**. Les sections suivantes (voir 1.6.4.2 et au-delà) approfondiront comment cette logique d’auto-organisation enrichit la **recommandation** par la capacité à intégrer de nouvelles entités et à effectuer une fusion dynamique des préférences et des contenus.

1.6.4.2. Mise à Jour Adaptative et Évolution dans le Temps

Dans un **système de recommandation**, la **distribution** évolue continuellement avec l’arrivée de nouveaux utilisateurs et contenus (articles, livres, films récents). Les préférences se modifient en fonction des tendances, des promotions ou des saisons. Les algorithmes classiques de filtering (factorisation, deep learning supervisé) doivent alors réestimer leurs paramètres, au risque d’exiger une nouvelle phase de réapprentissage ou un raffinement partiel, avec le danger de dérégler l’équilibre déjà atteint.

Le **Deep Synergy Learning (DSL)** adopte une démarche plus organique. Lorsque des entités inédites $\mathcal{E}_{\text{nouveauUser}}$ ou $\mathcal{E}_{\text{nouveauContenu}}$ apparaissent, on les **insère** directement dans le **Synergistic Connection Network**. Le mécanisme de mise à jour $\omega_{i,j}$ (section 1.4.5) se charge alors de **consolider** ou non les liaisons selon la **synergie** qu’elles dégagent. Ainsi, un nouveau contenu sera rapidement lié (pondérations positives) aux utilisateurs qui partagent des goûts proches ou qui interagissent avec ce contenu, conduisant à la formation d’un **cluster**. À l’inverse, un contenu ancien qui perd de l’intérêt voit ses pondérations $\omega_{u, \text{contenu}}$ décroître naturellement, jusqu’à s’isoler si plus personne ne s’y intéresse.

De plus, le **DSL** n’est pas cantonné aux seules entités “utilisateur” ou “contenu”. Des **entités contexte** comme la **saison** ($\mathcal{E}_{\text{été}}$), la **localisation** ($\mathcal{E}_{\text{Paris}}$), un **événement** ($\mathcal{E}_{\text{Noël}}$) peuvent influencer la **synergie**. Si la proximité entre un événement (Noël), certains contenus (articles cadeaux, recettes festives) et un profil d’utilisateurs se concrétise, alors les liens $\omega_{\text{Noël}, \text{contenu}}$ et $\omega_{\text{Noël}, \text{utilisateur}}$

s’amplifient, façonnant un **cluster** “Noël + X utilisateurs + Y contenus”. Lorsque la période de Noël s’achève, ces mêmes liaisons perdent peu à peu de leur attrait et se **délient**, laissant place à d’autres configurations contextuelles (telles que “soldes de janvier”). Le réseau se **reconfigure** ainsi localement, sans imposer un réentraînement massif ni une réinitialisation de tout le modèle, et garantit une adaptation continue au gré des changements de **distribution**.

1.6.4.3. Décision et Recommandation Basées sur les Pondérations Synergiques

Une fois le **réseau** de synergie $\{\omega_{i,j}\}$ formé ou relativement stabilisé (au moins à un instant donné), il devient possible d’en **extraire** un score de recommandation pour un **utilisateur** \mathcal{E}_u et un **contenu** \mathcal{E}_c . Une formule concevable, inspirée d’une multiplication matricielle partielle, consiste à sommer l’**influence** indirecte par d’autres entités $\{\mathcal{E}_k\}$ (utilisateurs proches, tags contextuels, règles). On peut ainsi écrire :

$$\text{score}_{u,c} = \sum_k \omega_{u,k} \omega_{k,c}.$$

Intuitivement, si l’utilisateur \mathcal{E}_u et un contenu \mathcal{E}_c ne partagent pas de lien direct, on évalue leur **affinité** en passant par des entités \mathcal{E}_k qui peuvent être d’autres **utilisateurs** (similaires) ou des **concepts** (genre musical, catégorie d’article). Cet agrégat de contributions évoque la structure de $(W^2)_{u,c}$ dans une multiplication de la matrice des pondérations W par elle-même, tout en permettant des **variantes** plus ajustées. Ces adaptations peuvent favoriser les **chemins pertinents** ou exploiter un **cluster** local pour améliorer la recommandation.

Lorsqu’on inclut des **entités symboliques** (section 1.5.7) représentant des **règles** ou **contraintes** métier, comme “Ne pas recommander les contenus PG-13 à un compte junior”, la **synergie** s’ajuste en conséquence. Si la liaison $\omega_{\text{rule}, \text{contenu}}$ signale une incompatibilité, elle diminue la pondération globale $\omega_{\text{user}, \text{contenu}}$, empêchant ainsi des recommandations inappropriées. Il en résulte un **réseau** qui intègre à la fois les **préférences** non supervisées (co-occurrences utilisateurs/contenus) et des **limites** ou **politique** explicites (règles symboliques). Le **SCN** en tient compte localement, ce qui équivaut à un **filtrage** ou à un **re-rank** dynamique, plus flexible qu’un pipeline figé. La **recommandation** se rapproche alors d’une conclusion collective de l’ensemble des entités, chaque liaison ω traduisant la **valeur** ou la **compatibilité** mutuelle.

1.6.4.4. Illustrations Concrètes

Plusieurs exemples concrets permettent de mieux percevoir la **souplesse** et la **richesse** du **Deep Synergy Learning (DSL)** appliqué au domaine des systèmes de recommandation ou, plus largement, à l’**aide à la décision** sur des ensembles utilisateur–contenu–contexte.

Dans une **plateforme e-commerce**, on dispose d’entités représentant les **utilisateurs** (avec leur historique et profil), les **produits** (catégories, marques, etc.), et d’éventuelles entités de **contexte** (temps fort comme soldes, promotions, événements saisonniers). Au lieu de programmer explicitement des algorithmes de segmentation, le **DSL** laisse ces entités s’**agréger** en **clusters** si elles constatent un **gain** mutuel. Un sous-groupe d’utilisateurs peut ainsi s’avérer très intéressé par

des produits d'une certaine marque ou style, formant un **cluster** qui se renforce dès lors que la synergie (p. ex. co-occurrence d'achats) s'avère constante. Les thématiques détectées évoluent naturellement en fonction des tendances ou de l'arrivée de nouvelles gammes de produits.

Dans un **service de streaming** (musique, vidéo), des entités peuvent correspondre aux **albums**, aux **artistes**, aux **playlists**, aux **tags** de genres (rock, pop, jazz), ainsi qu'aux **communautés** d'auditeurs. Le mécanisme de mise à jour $\omega_{i,j}$ valorise la proximité entre un auditeur et un certain artiste, ou entre un artiste et un tag de genre, etc. Un **cluster** cohérent peut alors émerger autour de l'auditoire d'une tendance musicale particulière (rock), intégrant les artistes, les playlists populaires associées et les auditeurs fidèles. Si un nouvel artiste rejoint la plateforme, on ajoute une entité $\mathcal{E}_{\text{artisteNew}}$; si celle-ci partage des caractéristiques avec le cluster rock (timbre de voix, style, tags), la **synergie** se trouve élevée, et l'artiste s'intègre au groupe. Les **recommandations** pour un auditeur donné s'en déduisent en analysant $\omega_{X,\text{artiste}}$ ou bien $(W^2)_{X,\text{artiste}}$ (cf. section 1.6.4.3), reflétant l'influence indirecte par d'autres auditeurs ou tags.

Dans un **système d'aide à la décision** industriel, on peut imaginer des entités décrivant différents **scénarios** (configurations de machines, paramètres de production), des **résultats passés** (qualité, temps de cycle), ainsi que des entités relatives aux **coûts** ou aux **ressources** (personnel, stocks). Les **liaisons** $\omega_{i,j}$ se renforcent lorsqu'une **combinaison** de paramètres s'avère fructueuse pour la production, ou que deux scénarios se révèlent similaires et conduisent à un résultat positif. Le **DSL** construit une **cartographie dynamique** de l'espace décisionnel. De nouveaux scénarios peuvent s'intégrer sans nécessiter de **réapprentissage global**, tandis que l'**auto-organisation** ajuste les liaisons si les conditions évoluent, comme un changement de cadence ou une modification des matières premières. Le tout offre une vision plus **distribuée** et **réactive** que des approches rigides, car les entités ne cessent de mettre à jour leurs liaisons pour refléter la **synergie** détectée.

1.6.4.5. Avantages et Défis

Dans un **système de recommandation** traditionnel, la construction d'un **pipeline** rigide est largement répandue, comme le collaborative filtering avec factorisation de matrice ; on effectue alors une estimation statique ou périodique, puis on l'emploie pour recommander. Le **Deep Synergy Learning (DSL)** adopte une **auto-organisation dynamique**. Lorsque les préférences d'un utilisateur évoluent ou qu'un contenu disparaît, les **connexions** ω s'ajustent localement, sans nécessiter de **recalcul global**. Cette **souplesse** améliore la réactivité à des évolutions telles que les promotions, la saisonnalité ou l'arrivée de nouveaux items.

Le **DSL** assure une robustesse et une parsimonie dans l'évolution du réseau. Les entités qui deviennent inactives, comme des utilisateurs peu actifs ou des contenus obsolètes, voient leurs liaisons ω diminuer naturellement lorsque leur synergie reste trop faible. Elles cessent alors d'interférer avec les autres recommandations, ce qui réduit le bruit et prévient la surcharge computationnelle. Un défi survient toutefois lorsque le nombre d'entités augmente considérablement. Il devient nécessaire d'adopter des heuristiques telles que la sparsification ou l'échantillonnage afin de garantir la scalabilité du réseau et d'éviter une explosion du nombre de connexions.

Le **DSL** permet également une évolution contextuelle fluide. Une nouvelle entité comme "festival d'automne", "temps pluvieux" ou "Black Friday" peut être introduite pour ajuster dynamiquement

la synergie entre les éléments du réseau. Ces entités contextuelles influencent directement les connexions ω et révèlent quels utilisateurs et contenus sont concernés par un événement donné. Aucun réapprentissage massif n'est requis. Cela permet une adaptation continue où les connexions évoluent en temps réel selon les événements, sans imposer de rigidité structurelle.

En dernier lieu, l'**explicabilité** bénéficie d'une amélioration notable. Plutôt que de s'appuyer sur des facteurs latents abstraits difficiles à interpréter, un **cluster** émergent permet une compréhension plus directe. Il regroupe un ensemble d'utilisateurs avec des contenus spécifiques, révélant une structure cohérente et identifiable. Par exemple, une entité représentant le genre musical "jazz" peut se lier naturellement avec une entité "club de blues" et un sous-groupe d'auditeurs partageant ces préférences. Cela offre une explication claire et justifiable. Lorsqu'un utilisateur reçoit une recommandation, il devient possible de formuler une justification explicite comme "vous appartenez au cluster jazz-blues, ce qui explique cette suggestion". Cette approche dépasse la factorisation matricielle classique, apportant une transparence accrue et facilitant l'interprétation des décisions prises par le système.

1.6.5. Surveillance, Diagnostic Médical et Anomalies

Les principes d'**auto-organisation** et de **coopération adaptative** propres au **Deep Synergy Learning (DSL)** ne se cantonnent pas aux domaines précédemment cités (vision, audio, robotique, recommandation). Ils trouvent également une place toute particulière dans les **applications de surveillance**, de **diagnostic**, et de **détection d'anomalies**, qu'il s'agisse de contexte médical ou industriel. Dans ces environnements, on gère souvent d'importants **flux de données** en temps réel, provenant de **capteurs variés** (biologiques, physiques) qu'il faut **coordonner** pour repérer des situations anormales ou des symptômes précoces. Les systèmes classiques reposent sur des algorithmes d'**anomaly detection** (souvent supervisés, semi-supervisés, ou basés sur des seuils statiques) qui peuvent se révéler rigides ou trop dépendants d'un jeu de labels.

Le **DSL** apporte une **flexibilité** en structurant des **entités d'information** représentant divers signaux comme des capteurs, des paramètres cliniques ou des logs machines. Ces entités sont reliées par des **pondérations synergiques** qui évoluent au fil du temps, assurant une **auto-organisation** et une adaptation continue aux variations des données. Ainsi, l'**auto-organisation** permet de détecter des groupes d'observations cohérentes et, réciproquement, d'identifier des entités qui s'**écartent** de la synergie générale. Cette capacité de **repérage** local et de **réaction** dynamique rend le **DSL** adapté aux systèmes de **surveillance**, où des capteurs ou des caméras isolent des comportements inhabituels. Il s'avère aussi pertinent pour les **diagnostics médicaux**, en intégrant imagerie, analyses biologiques et relevés de signes vitaux pour identifier des pathologies. Enfin, dans la **maintenance prédictive**, il permet de détecter les anomalies en observant les écarts de certains capteurs par rapport au comportement attendu du **cluster**.

Nous allons voir plus en détail comment le DSL prend en compte la variété des signaux et repère, via la synergie (ou son absence), ce qui sort du lot et pourrait constituer une alerte ou un symptôme.

1.6.5.1. Systèmes de Surveillance : Entités Multi-Capteurs, Scènes Vidéo

Dans de nombreux dispositifs de **surveillance vidéo**, les algorithmes classiques (souvent un pipeline à base de CNN et de suivi d'objets) doivent gérer un **flot** de données considérable et détecter en continu tout comportement ou événement anormal. Cette approche peut se révéler fragile face à la complexité des scènes ou à la diversité des conditions. Le **Deep Synergy Learning (DSL)** propose une alternative plus **auto-organisée** et **réactive**, en représentant chaque composante (segment visuel, contexte, schéma habituel) sous forme d'entités et en laissant la **synergie** orienter la structuration.

Dans le cadre d'une **vidéo** de surveillance, on peut définir des entités $\mathcal{E}_{\text{seg},k}$ correspondant à différents **blobs** ou **segments** détectés dans la scène (régions de mouvement, silhouettes, etc.). D'autres entités $\mathcal{E}_{\text{contexte}}$ décrivent, par exemple, l'heure, la zone surveillée, les règles de fréquentation habituelles. Enfin, des entités $\mathcal{E}_{\text{pattern}}$ peuvent caractériser des **patrons** récurrents (le nombre moyen de personnes, la vitesse de déplacement habituelle, les régions autorisées ou interdites). Au sein du **Synergistic Connection Network**, les liaisons ω évoluent en fonction de la **cohérence** détectée. Si un segment vidéo "blob A" suit un déplacement conforme aux attentes en termes de vitesse et d'emplacement, la synergie reste élevée. Lorsqu'un comportement s'écarte de l'ordinaire (un mouvement inhabituel, une densité de blobs anormale, etc.), la **pondération** $\omega_{\text{seg}, \text{pattern}}$ chute, isolant le segment considéré et signalant une anomalie. L'algorithme peut alors émettre une alarme ou enclencher une inspection spécifique, sans nécessiter de recompiler ou de réentraîner un détecteur complet.

Dans la **surveillance** élargie, d'autres capteurs peuvent être intégrés, comme les détecteurs de mouvement (PIR), les capteurs de bruit, de température ou de fumée. Le **DSL** traite aisément cette hétérogénéité en insérant des **entités** $\mathcal{E}_{\text{mouvement}}$, \mathcal{E}_{son} , $\mathcal{E}_{\text{température}}$, dont les synergies s'**auto-régulent** avec les entités visuelles. Si un choc violent est détecté par \mathcal{E}_{son} alors que la vidéo ne rapporte aucun mouvement, la liaison $\omega_{\text{son}, \text{cam}}$ diminue, suggérant un capteur défaillant ou un phénomène localisé sans trace visuelle. À l'inverse, la convergence de plusieurs capteurs (caméra détectant flammes, capteur son captant un "boom", etc.) renforce les liaisons, formant un **cluster** d'alerte qui intègre différentes sources. Cette **auto-organisation** renforce la **résilience** du système. Les signaux incertains sont naturellement atténués, tandis que les corrélations significatives entre capteurs entraînent une consolidation des liaisons, améliorant ainsi la détection et l'identification des événements critiques.

1.6.5.2. Diagnostic Médical : Entités Physiologiques et Multimodalité



Dans un **contexte médical**, les données proviennent de multiples **sources**. On dispose de l'**imagerie** (IRM, rayons X, scanner), de **tests biologiques** (formule sanguine, biochimie), de **signes vitaux** (ECG, pression, saturation en O₂) et de **symptômes** rapportés, voire de données **génétiques** plus complexes. Un **réseau** de neurones classique, comme un CNN spécialisé dans l'interprétation d'IRM, a tendance à se focaliser sur un **type** de signal. Or, certains diagnostics difficiles (cancers rares, syndromes complexes) exigent un **croisement** entre diverses modalités, chacune partielle, parfois bruitée.

Le **Deep Synergy Learning (DSL)** se conçoit alors comme un **réseau** auto-organisé où des **entités** $\{\mathcal{E}_{\text{irm}}, \mathcal{E}_{\text{analyseSang}}, \mathcal{E}_{\text{ecg}}, \dots\}$ représentent chacune un bloc de données ou un test particulier. Le **score** de synergie $S(\mathcal{E}_i, \mathcal{E}_j)$ indique si deux sources convergent vers une suspicion commune (par exemple, des anomalies visibles à l'IRM qui correspondent à des marqueurs sanguins inhabituels). Les **pondérations** $\omega_{i,j}$ se renforcent lorsque la coopération de deux entités apporte un **gain** (détection corrélée), et un **cluster** pathologique peut se constituer, regroupant un sous-ensemble de paramètres (imagerie, analyses, symptômes) indiquant une même pathologie ou un syndrome précis.

L'évolution **temporelle** des données médicales s'adapte naturellement aux fluctuations de l'état du patient. Lorsqu'un paramètre revient à la normale, la synergie avec les marqueurs pathologiques s'affaiblit, entraînant une **désolidarisation** progressive du **cluster** associé à la maladie. À l'inverse, un nouveau symptôme peut s'y **agréger** dès qu'il présente une similarité ou une co-occurrence marquée avec d'autres facteurs pathologiques existants. Le **diagnostic** évolue en continu, ajustant la structure du réseau selon l'**état** du patient.

Cette dynamique peut être renforcée par un **signal** de validation clinique. Lorsqu'un diagnostic s'avère correct, les liens responsables de cette détection se consolident, améliorant la robustesse du système. Ce mode d'**auto-organisation** se distingue des pipelines rigides, qui traitent chaque modalité indépendamment (imagerie, analyses biologiques). En intégrant **toutes** les sources de données, les pondérations ω révèlent la **coopération** la plus pertinente à un instant donné, optimisant la prise de décision.

1.6.5.3. Détection d'Anomalies Industrielles : Auto-Organisation et Isolement

En milieu **industriel** (production, logistique, transports), la **maintenance prédictive** implique une surveillance continue des capteurs (température, vibrations, courants moteurs, etc.) afin de détecter des signaux précurseurs de panne. Dans un **Deep Synergy Learning (DSL)**, on peut représenter chaque **capteur** par une entité $\mathcal{E}_{\text{capteur}}$, ajouter des entités décrivant des **normes** ou **règles** de fonctionnement (certaines pouvant être symboliques, voir section 1.5.7), ainsi que des **entités contextuelles** (charge de travail, température ambiante, etc.). Ces entités s'**auto-organisent** et forment un **cluster** principal (ou plusieurs) lorsqu'elles partagent un comportement cohérent.

Une **anomalie** se manifeste lorsque la **synergie** $S(\mathcal{E}_k, \mathcal{E}_{\text{ref}})$ entre un capteur \mathcal{E}_k et le **cluster** de référence (normes, capteurs stables) chute. Contrairement à un système à **seuils fixes** appliqués indépendamment à chaque capteur, le **DSL** évalue chaque capteur par rapport à l'**ensemble** du réseau. Si un capteur **diverge** seul et ne suit plus la synergie habituelle, ses pondérations ω diminuent, l'isolant progressivement et signalant une **anomalie** potentielle. À l'inverse, si plusieurs capteurs entament des dérives **cohérentes**, ils peuvent former un **sous-cluster** spécifique (un "cluster d'anomalie"), laissant deviner un défaut commun (ex. un roulement en fin de vie qui génère bruit et surchauffe).

Au fil du temps, un capteur fautif peut s'**isoler** progressivement si ses pondérations ω_k diminuent, indiquant une perte de synergie avec le reste du réseau. Il peut aussi entrer en **coopération** avec d'autres entités dérivantes, formant un **cluster** d'anomalie plus large. Si cette situation persiste et que la synergie interne de l'anomalie (par exemple, bruit anormal, vibrations, température de roulement) continue d'augmenter, cela peut déclencher un **signal** de maintenance, suggérant une panne imminente.

Cette **auto-organisation** rend la détection plus **adaptative**. Le **réseau** ajuste progressivement les **liaisons** en renforçant celles qui restent cohérentes et en réduisant celles qui s'écartent, signalant ainsi des anomalies potentielles. On évite ainsi les contraintes d'un pipeline rigide ou de seuils par capteur, car la logique de synergie estime la **relation** du capteur à l'ensemble du **cluster** de fonctionnement normal, autorisant un repérage de dérive **contextualisé** et évolutif.

1.6.5.4. Avantages et Défis : la Logique du DSL en Alerte et Diagnostic

Les applications de **Deep Synergy Learning (DSL)** dans le domaine de la **surveillance**, de la **maintenance** ou du **diagnostic** s'appuient sur sa capacité à constituer des **clusters** de manière autonome, en agrégeant des **entités** multiples (capteurs, tests, imagerie, etc.). Contrairement aux solutions qui demandent une fusion de flux orchestrée manuellement, le **DSL** laisse les pondérations $\omega_{i,j}$ évoluer en fonction de la **synergie** locale, formant des sous-groupes pertinents sans pipeline imposé.

Cette approche confère une **flexibilité** face aux données partielles ou intermittentes. Des capteurs sujets à des interruptions ou à des pannes peuvent se retirer du **cluster** principal dès lors que leur **synergie** avec les autres entités s'effondre, puis se réinsérer lorsque leur fiabilité s'améliore. On évite ainsi qu'un capteur défaillant ne paralyse l'ensemble.

La **détection de situations inédites** (anomalies, configurations nouvelles) bénéficie particulièrement de la logique non supervisée du **DSL**. Quand un groupe d'entités dérive au regard de la norme, il se regroupe spontanément en **cluster** d'anomalie, sans que l'on ait à pré-définir les classes d'incidents. Cela autorise un repérage de schémas insoupçonnés ou de pannes rares.

Les mécanismes d'**adaptation continue** permettent une réactivité en contexte dynamique. Les pondérations ω se réajustent dès qu'un changement est détecté, que ce soit une évolution de l'état du patient en milieu médical ou une transition vers un nouveau régime en milieu industriel. Ce fonctionnement n'exige pas un **réapprentissage** exhaustif à chaque changement.

Cependant, le **principal défi** réside dans la **taille** du graphe et la croissance exponentielle du nombre de liens $\{\omega_{i,j}\}$ quand les entités se multiplient (nombre élevé de patients, de capteurs, de variables, etc.). Il convient d'appliquer des mesures de **parsimonie** (un seuil en-deçà duquel on supprime des connexions), des procédures de **sparsification** dynamique, ou encore des **misés à jour** localisées (échantillonnage partiel, focus sur un sous-ensemble d'entités), pour maintenir la **scalabilité** et la réactivité du système.

1.6.6. Planification et Optimisation dans l'Industrie 4.0

Les secteurs industriels intègrent une **connectivité** croissante, où machines, capteurs, robots, stocks et systèmes de gestion interagissent en temps réel. Cette dynamique incarne la vision de l'**Industrie 4.0**, où l'optimisation repose sur l'échange fluide de données et l'**auto-adaptation** des processus. Les défis qui en découlent — **planification** distribuée, **optimisation** des flux, **maintenance prédictive** et **gestion adaptative** — peuvent se révéler extrêmement complexes. Les méthodes traditionnelles (ordonnancement déterministe, programmation linéaire classique, heuristiques fixes, etc.) montrent souvent leurs limites face à la **dynamique** croissante (marchés volatils, nouvelles commandes urgentes, pannes inattendues). Dans ce cadre, le **Deep Synergy Learning (DSL)**, avec son **mécanisme d'auto-organisation** et de **synergie adaptative**, ouvre de nouvelles pistes pour réaliser une **planification plus flexible** et des **schémas d'optimisation** capables de se réajuster en continu.

Dans cette section (1.6.6), nous examinons :

- Les raisons pour lesquelles l'**Industrie 4.0** exige un modèle **évolutif** et **distribué**,
- Comment le **DSL** modélise des entités (machines, stocks, commandes, flux de transport) reliées par des **pondérations synergiques**,
- De quelle manière ces liens évoluent pour **planifier** et **optimiser** les ressources en temps réel,
- Des exemples concrets où le **DSL** peut surpasser un pipeline d'optimisation statique, grâce à sa **dynamique d'adaptation** et à sa capacité de **clustering** auto-organisé.

1.6.6.1. Les Défis Industriels dans un Environnement 4.0

Dans l'Industrie 4.0, on fait face à une **abondance** de composants interdépendants. Les **machines** possèdent chacune un état (vitesse, disponibilité, pannes potentielles), tandis que les **chaînes de production** sont subdivisées (stations, lignes, buffers, flux de pièces). Les **stocks** (pièces détachées, semi-finis, produits finis) changent au fil des demandes, et le **transport** peut impliquer différents véhicules (AGV, drones, camions). Enfin, les **ordres de fabrication** évoluent dès lors que des clients passent de nouvelles commandes ou modifient leur planning.

Dans les approches classiques d'**ordonnancement** ou de **planification** (flow shop, job shop, heuristiques variées), on définit un **problème combinatoire** qu'on tente de résoudre sous des **hypothèses** de stabilité. En pratique, l'environnement évolue en permanence. Les machines peuvent tomber en **panne**, un retard logistique peut compromettre la disponibilité d'une pièce critique, de nouvelles **commandes** peuvent arriver de manière imprévue, et la **capacité** des machines peut fluctuer en raison de maintenances ou de changements d'outillage. Chaque perturbation exige souvent un **réajustement** conséquent, voire une **réinitialisation** complète d'un algorithme d'ordonnancement.

Le **Deep Synergy Learning (DSL)** propose un **changement de perspective**. Au lieu de modéliser un ordonnancement global, on conçoit chaque **composant** — machine, ordre, stock, mode de transport — comme une **entité** \mathcal{E}_i . Les **liaisons** $\omega_{i,j}(t)$ traduisent la **coopération** ou la **complémentarité** entre entités. Par exemple, deux machines $\mathcal{E}_{\text{machineA}}$ et $\mathcal{E}_{\text{machineB}}$ voient leur lien s'intensifier si elles se passent fréquemment des pièces sans blocage, ou si l'on constate un gain productif à enchaîner les opérations A puis B. Un **stock** $\mathcal{E}_{\text{stock}}$ et un **transport** \mathcal{E}_{AGV} coopèrent s'ils se synchronisent efficacement pour déplacer des lots. De même, un **ordre** $\mathcal{E}_{\text{ordre}}$ accroît sa synergie avec une machine spécifique si cette dernière exécute la tâche en temps et qualité optimales.

Lorsqu'un **incident** surgit (panne machine, retard de livraison, surcharge de commande), la **dynamique d'auto-organisation** du DSL réévalue les liaisons ω . Un composant en panne voit ses liens se dégrader, réduisant son impact dans le flux. Inversement, si une nouvelle machine arrive ou si un transport reprend sa disponibilité, on insère une **entité** inédite dans le **SCN**, laissant la synergie se construire là où un **gain** apparaît. Cette **reconfiguration** s'effectue **localement**, évitant de relancer un algorithme d'ordonnancement global. Le **réseau** s'ajuste de manière "vivante", reflétant la réalité mouvante de l'Industrie 4.0.

1.6.6.2. Représentation des Flux et Clusters Auto-Organisés

Dans une optique d'**Industrie 4.0**, la mesure de **synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$ entre deux entités \mathcal{E}_i et \mathcal{E}_j peut prendre plusieurs formes. Les machines, les stocks ou les ordres de production entretiennent par exemple un **taux de transfert** élevé lorsque la première fournit régulièrement des pièces à la seconde sans créer de file d'attente. Un **ordre** de production se révèle quant à lui **complémentaire** d'un stock donné si cet ordre requiert précisément les ressources contenues dans ce stock. Par ailleurs, la **performance** d'une association machine–ordre s'accroît lorsque le couple aboutit à un temps de cycle réduit et à un rendement supérieur. Lorsque la **synergie** apparaît remarquable entre

deux entités, la pondération reliant ces entités se trouve renforcée (voir section 1.4.5 sur la mise à jour locale des pondérations).

À mesure que certaines combinaisons s'avèrent profitables (machine–stock–ordre–transport, par exemple), leur coopération se **solidifie**, faisant émerger des **clusters** ou sous-réseaux. Chaque **cluster** correspond alors à une configuration de production considérée stable ou efficace, puisqu'il regroupe des entités dont les pondérations mutuelles ω sont élevées. On observe ainsi la formation d'un **cluster** regroupant {Machine1, StockA, TransportX, OrdreB}, car ces éléments s'articulent en un **flux** cohérent. La Machine 1 puise ses ressources dans le Stock A, le Transport X assure la logistique en temps réel, et l'Ordre B bénéficie d'un traitement optimisé.

Dans le cadre d'une planification traditionnelle, on conçoit des **lignes** ou **cellules** de production de manière analytique. Le **Deep Synergy Learning (DSL)** repose sur une **auto-organisation dynamique** où les **liaisons** se renforcent lorsque des machines, des stocks ou des modules de transport démontrent une **compatibilité fonctionnelle**. Cette dynamique favorise l'émergence **spontanée** de **clusters** optimisés, sans nécessiter une configuration prédéfinie du système. Quand un composant (machine, stock ou transport) **dysfonctionne** ou s'avère insuffisamment utile, la **synergie** avec le reste du réseau chute et le composant se trouve mis à l'écart. De nouveaux liens peuvent alors se former vers des entités de **remplacement**, réorientant la production de façon plus souple qu'un pipeline rigide. Ce phénomène procure à l'ensemble une **plasticité** qui évoque un système "**vivant**", réactif aux perturbations et aux changements d'outillage ou de répartition des tâches, conformément à l'esprit de l'Industrie 4.0 décrite en section 1.6.6.1.

1.6.6.3. Approche Dynamique : Adaptation en Temps Réel

Dans un environnement **Industrie 4.0**, le système gère un **flux continu** d'événements tels que nouvelles commandes, variations de stocks, maintenances ou retards logistiques. Le **DSL** ajuste en **temps réel** les **liaisons synergiques** entre machines, ressources et transports. Les connexions utiles se renforcent tandis que les moins pertinentes s'affaiblissent, assurant une **auto-organisation dynamique** sans réajustements manuels. Les algorithmes de planification classiques (règle de Johnson, heuristiques de scheduling, etc.) se montrent performants pour un **problème** donné (flow shop, job shop), mais perdent en **flexibilité** lorsqu'on modifie le problème à la volée, car il faut alors réinitialiser ou réactualiser le schéma de planification.

Au **Deep Synergy Learning (DSL)**, la **dynamique** d'auto-organisation se renouvelle en continu. Les **pondérations** $\omega_{i,j}(t)$ reliant des entités (machines, ordres, stocks, transports...) se remanient itération après itération en fonction de la **synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$. Lorsqu'un **ordre urgent** survient, le système introduit une entité $\mathcal{E}_{\text{ordreUrgent}}$ dans le **Synergistic Connection Network**, et celle-ci voit ses **liaisons** $\omega_{\text{ordreUrgent},k}$ se renforcer rapidement dès que l'on identifie un gain potentiel avec des machines ou des transporteurs capables de le traiter à temps. Si, au contraire, un transport $\mathcal{E}_{\text{transport}}$ se montre sous-dimensionné (accumulation de pièces, délais excessifs), sa **synergie** avec le reste du réseau chute, ce qui oriente la production ou la logistique vers d'autres entités plus aptes à satisfaire la demande.

Le **DSL** ne cherche pas à figer une **solution** optimale à un instant donné comme un algorithme de scheduling. Il maintient un **réseau dynamique**, où les **clusters** de machines, ordres et stocks

s'adaptent aux événements. En intégrant un **signal global** de performance comme le throughput, le temps de cycle ou le taux de satisfaction client, le système ajuste la **synergie** de manière dynamique. Les entités impliquées dans une configuration efficace renforcent leurs **liaisons** ω , tandis que celles associées à une baisse de performance voient leurs connexions s'affaiblir. Cette adaptation continue permet une **auto-organisation** où chaque entité ajuste sa coopération localement, sans nécessiter une refonte complète de la planification à chaque nouvel incident.

1.6.6.4. Cas d'Exemple : Usine Modulable

Dans une **usine** où les postes sont **reconfigurables** et la logistique assurée par des **AGV** (véhicules autonomes), les produits fabriqués peuvent changer selon la **demande**. Le **Deep Synergy Learning (DSL)** s'avère alors bénéfique pour une **gestion** plus flexible, en remplaçant ou en complétant les méthodes de planification traditionnelles. On peut définir un **réseau** d'entités $\{\mathcal{E}_{\text{Poste1}}, \mathcal{E}_{\text{Poste2}}, \dots, \mathcal{E}_{\text{AGV1}}, \mathcal{E}_{\text{AGV2}}, \dots, \mathcal{E}_{\text{Stock1}}, \dots\}$ et des entités "ProduitA", "ProduitB" (avec la possibilité d'en ajouter de nouveaux). Chaque poste ou AGV entretient des **liaisons** $\omega_{i,j}$ qui se **renforcent** ou se **délient** selon la **synergie** constatée. Si un **poste** particulier traite un certain **produit** avec un taux de réussite élevé, la pondération $\omega_{\text{Poste}, \text{Produit}}$ augmente. De même, un **AGV** se connecte aux postes ou stocks avec lesquels il collabore efficacement, formant ainsi un **cluster** $\{\text{AGV1}, \text{Poste1}, \text{Poste2}, \text{Stock}\}$ si ce groupe de composants assure un flux fluide de pièces.

Lorsqu'on introduit un **nouveau** produit (par exemple "ProduitC"), la synergie se construit **naturellement** autour des postes adéquats. Le **SCN** évolue sans qu'il soit indispensable de reconfigurer entièrement le pipeline ou de lancer un **solveur** d'ordonnancement intégral. Les entités reliées à l'ancien produit peuvent **conserver** leurs liens, tandis que les postes ou AGV inadaptés à "ProduitC" voient leur pondération chuter et se déconnectent graduellement du nouveau cluster. Cette **dynamique** garantit une **réactivité** efficace. En cas de **panne** sur Poste1, la pondération $\omega_{\text{AGV}, \text{Poste1}}$ diminue, incitant l'AGV à se rediriger vers un autre poste, comme $\mathcal{E}_{\text{Poste3}}$. La **synergie** avec ce nouveau poste augmente alors naturellement, optimisant ainsi l'adaptation du réseau.

Ce **mécanisme d'auto-organisation** continue favorise la **tolérance** aux pannes et l'adaptation à l'introduction de nouvelles gammes de produits. Les entités inopérantes se "détachent" naturellement (perdent leurs liaisons), et les nouvelles entités intéressantes (nouveaux postes, nouveaux produits) trouvent leur place dans le **Synergistic Connection Network** au fil de la **dynamique** des pondérations. Cela évite la nécessité de **recalculer** périodiquement un ordonnancement global ou de procéder à des heuristiques lourdes, tout en exploitant les **principes** du DSL (sections 1.4.3 et 1.5.4) pour maintenir la cohérence globale du système.

1.6.6.5. Limites et Évolutions : Complexité, Parcimonie, Hybridation

Dans un **environnement industriel** où l'on introduit sans cesse de nouvelles entités (machines, stocks, ordres...), la **taille** du graphe $\{\omega_{i,j}\}$ peut rapidement devenir prohibitive. Pour préserver la faisabilité, il est essentiel d'envisager des stratégies de **parcimonie** (seuils, coupe des liens trop faibles) ou des **algorithmes locaux** où seules les pondérations pertinentes se mettent à jour, évitant ainsi un recalcul exhaustif à chaque itération. Des **contraintes** peuvent également exclure certains

liens d'emblée (une machine non qualifiée pour un ordre, ou un transport incapable de prendre en charge un certain poids), imposant une **pénalisation** ou un **masquage** dans la fonction de synergie.

De plus, lorsque des **règles logiques** (contraintes de sécurité, de qualité) s'ajoutent, on peut intégrer des **entités symboliques** (cf. section 1.5.7). Le **DSL** considère alors la **coopération** entre ces entités symboliques et les entités opérationnelles (machines, ordres...), de sorte que la formation d'un **cluster** illégal (par exemple, un groupement violant une norme de sécurité) n'aboutisse jamais, faute de synergie. Cette hybridation **neuro-symbolique** enrichit la dynamique locale (ω) tout en respectant les cadres réglementaires ou métier.

Il est par ailleurs envisageable de **combiner** le **DSL** à des heuristiques ou solveurs d'ordonnancement établis. Une possibilité consiste à laisser l'**auto-organisation** déterminer un sous-ensemble de ressources potentiellement efficaces, puis à faire appel à un **solveur** classique pour finaliser le planning sur ce sous-ensemble, réduisant la dimension du problème. À l'inverse, un solveur peut fournir une solution initiale, que le **DSL** ajuste localement au gré des aléas (pannes, nouvelles commandes). Cette **collaboration** exploite la force de l'**auto-organisation** (réactivité, plasticité) tout en bénéficiant de méthodes d'**optimisation** éprouvées pour parfaire le résultat. Ainsi, la **plasticité** du **DSL** et la **rigueur** de l'optimisation se conjuguent pour répondre aux exigences variées de l'Industrie 4.0.

1.6.7. Perspectives pour la Recherche Fondamentale en IA Forte

Les précédentes sous-sections (1.6.1 à 1.6.6) ont démontré la **portée pratique** du **Deep Synergy Learning (DSL)** dans divers domaines applicatifs (vision, audio, robotique, recommandation, diagnostic, etc.). Cependant, l'**ambition** du **DSL** ne se limite pas à une simple amélioration des performances ou de la flexibilité dans ces tâches spécialisées. À un niveau plus **fondamental**, l'approche synergique **ouvre des pistes** de recherche susceptibles de rapprocher l'IA de la notion d'**IA Forte** (ou **IA Généralisée**), c'est-à-dire d'une intelligence capable d'**apprendre** et de **raisonner** de manière autonome et générale, au-delà de cadres strictement définis. Cette section (1.6.7) explore plusieurs **perspectives** théoriques et conceptuelles. D'abord, le **DSL** et ses **mécanismes d'auto-organisation** présentent des similarités avec les **systèmes cognitifs**, qu'il s'agisse du cerveau ou d'écologies d'informations. Ensuite, il offre la possibilité d'**intégrer** à la fois des dimensions sub-symboliques, reposant sur l'apprentissage à partir de données massives, et des **aspects symboliques** ou logiques (voir 1.5.7). Par ailleurs, des **propriétés émergentes** se manifestent, telles que la formation de représentations, de micro-réseaux cognitifs et de processus d'auto-adaptation qui dépassent la simple exécution d'une tâche. Enfin, l'approche soulève plusieurs **défis mathématiques** et philosophiques liés à l'ambition d'une **IA Forte** fondée sur le **DSL**.

1.6.7.1. Au-delà de l'Apprentissage Supervisé : Vers l'Auto-Construction de la Connaissance

Une large part des avancées récentes en **IA** se fonde sur l'**optimisation globale** de réseaux neuronaux massifs par **rétropropagation**, qu'il s'agisse de données annotées en abondance ou d'un pré-entraînement auto-supervisé (approche de type GPT, BERT). Bien qu'efficaces, ces modèles conservent plusieurs caractéristiques. Ils reposent sur une **architecture** linéaire ou

faiblement hiérarchisée, même en présence de mécanismes comme les multi-têtes d'attention. Leur optimisation dépend d'une **descente de gradient** ou d'une procédure équivalente visant à minimiser un coût défini globalement. Enfin, leur **adaptabilité locale** reste restreinte. Lorsque la distribution des données change, un **réapprentissage** par fine-tuning devient souvent indispensable, ce qui expose le modèle au risque d'un **écrasement partiel** des connaissances antérieures, connu sous le nom de **catastrophic forgetting**.

Le **Deep Synergy Learning (DSL)** présente une **dynamique d'auto-organisation** plus proche de l'esprit de **systèmes complexes** (biologiques ou écologiques), où chaque **entité** (sous-module, neurone, flux d'information) gère ses **liaisons** suivant la **synergie** qu'il perçoit avec d'autres entités. Plutôt que de minimiser une **unique fonction de coût**, on autorise l'émergence de **clusters** et la **reconfiguration permanente** des connexions $\omega_{i,j}$. Cette approche "**bottom-up**" s'avère potentiellement féconde pour développer des **mécanismes cognitifs** plus avancés et moins tributaires d'un lourd apprentissage supervisé.

Dans une perspective d'**IA Forte**, on envisage souvent un agent qui **accumule** graduellement des connaissances sans les oublier, tout en **réinterprétant** ces savoirs au fil des expériences. Le **DSL** y répond en laissant chaque **entité** conserver sa mémoire locale (paramètres, historique) et en **réévaluant** les liaisons ω de façon continue. On obtient ainsi un processus **incrémental** où de nouvelles **entités** émergent, représentant des concepts ou des données. Les **liaisons** se **créent** ou se rompent en fonction de la **pertinence** et de la **coopération** mesurée. Les **clusters** de connaissances se **stabilisent** lorsque la synergie interne reste forte, ou se **scindent** si celle-ci décline. L'agent parvient ainsi à **auto-structurer** ses connaissances, à l'image d'un cerveau qui forme et dissout continuellement des **assemblées** neuronales en fonction des interactions et de l'expérience.

Cette organisation, plus **distribuée** et **évolutive**, pose les bases d'une **auto-construction** de la connaissance, où l'on n'est plus tributaire d'un unique label ou d'une unique fonction de perte, mais d'une dynamique **locale** entre les **entités** qui renforcent ou amoindrissent leurs liens au gré de leur coopérativité. Elle suggère la possibilité d'un **savoir** plus flexible et cumulatif, réduisant la dépendance à un entraînement exhaustif et favorisant une **adaptation** continue aux données nouvelles ou aux changements de contexte.

1.6.7.2. Vers une Cognition Distribuée et Émergente

Le **Deep Synergy Learning (DSL)** trouve ses racines (section 1.3.1) dans l'observation des **systèmes biologiques** exhibant de la **plasticité** (synaptique chez les neurones), ou des phénomènes d'**auto-organisation** (colonies d'insectes, tissus cellulaires). Dans la quête d'une **IA Forte**, un **réseau** d'entités peut servir de **substrat cognitif** en intégrant plusieurs dimensions.

Les **entités** \mathcal{E}_i regroupent aussi bien des *features* perceptuelles issues de flux sensoriels que des **concepts abstraits** ou des **modules symboliques** encapsulant des règles logiques.

Les **liaisons synergiques** $\omega_{i,j}$ se renforcent selon un mécanisme inspiré du renforcement "hebbien généralisé" dès lors qu'une **coopération** entre entités procure un **gain**, qu'il s'agisse de performance, de co-information ou de simplification.

Les **clusters** qui émergent forment des circuits cognitifs spécialisés, participant au traitement conceptuel, à la mémoire de travail, à la planification ou à l'intégration de signaux multiples.

Au sein de ce **substrat**, chaque **cluster** reflète potentiellement une micro-théorie ou un schéma conceptuel, regroupant des entités pour former une **hypothèse**, un **concept**, voire un **plan d'action**. La structuration évolue de manière **dynamique**. Les assemblées émergent ou se dissolvent en fonction de la synergie détectée, permettant une **reconfiguration continue** des connaissances et des idées. Contrairement aux architectures statiques à couches fixes, cette organisation s'**auto-régule**, équilibrant **plasticité** (création et suppression de liens) et **stabilité** (maintien des clusters pertinents).

En examinant la **dynamique** globale, on observe l'émergence de **processus cognitifs** spécifiques. Une **forme d'attention** se manifeste lorsqu'un sous-réseau capte et monopolise la synergie disponible, créant un **focus transitoire** qui oriente temporairement le traitement des informations. Des **boucles** de rétroaction positives ou négatives peuvent générer des **oscillations**, analogues à des rythmes cognitifs ou des alternances d'états mentaux. La **conservation** d'équilibres permet au système de préserver des organisations robustes, tout en se montrant apte à accueillir de nouvelles liaisons ou à rompre les moins pertinentes.

De tels phénomènes, absents d'un CNN ou d'un Transformer (qui restent malgré tout dans une structure de pipeline, bien que complexe), s'apparentent plutôt à un **processus** cognitif **distribué**, dans lequel l'information chemine entre assemblées modulaires qui se constituent et se dissolvent librement. On entrevoit là un **jalon** vers des agents en mesure de **construire** et de **réviser** en continu leurs hypothèses ou "théories", de manière plus proche de la **biologie** et des **neurosciences** que de la minimisation stricte d'une fonction de coût globale.

1.6.7.3. Intégration Symbolique : Socle pour le Raisonnement Abstrait

Dans une **démarche d'IA Forte**, il ne suffit pas d'analyser des signaux sub-symboliques comme les images, les capteurs ou les séquences temporelles ; il faut aussi être capable de manipuler des **représentations abstraites** et des **règles logiques complexes**. Les réseaux neuronaux profonds, bien qu'efficaces sur des volumes massifs de données, éprouvent des difficultés à exécuter un raisonnement explicitement symbolique, sauf si on leur adjoint des modules dédiés. Le **Deep Synergy Learning (DSL)** propose un **cadre unifié** où des entités **symboliques** (comme $\mathcal{E}_{\text{rule}}$, $\mathcal{E}_{\text{concept}}$) coexistent avec des entités **sub-symboliques** ($\mathcal{E}_{\text{sensor}}$, $\mathcal{E}_{\text{feature}}$), comme décrit à la section 1.5.7. Lorsqu'une **règle** s'applique à une **configuration perceptive**, la liaison $\omega_{\text{rule}, \text{sensor}}$ se renforce dès que la synergie entre elles se révèle fructueuse (information mutuelle élevée, gain de performance, amélioration de la précision). Il en résulte un **micro-réseau** où un ensemble de règles se coordonne avec un ensemble d'observations, sans qu'il soit nécessaire d'établir un cloisonnement strict entre logique et perception.

Dans cette perspective, la dynamique des pondérations ω ne se limite pas au rapprochement d'entités visuelles ou auditives. Elle permet également de **connecter des concepts abstraits à des indices sensoriels** confirmant leur validité ou, à l'inverse, d'affaiblir une règle qui ne s'accorde pas aux observations. Cela conduit à un **raisonnement logico-perceptif fluide**. Lorsqu'une hypothèse symbolique $\mathcal{E}_{\text{rule}}$ reçoit un soutien sensoriel de $\mathcal{E}_{\text{sensor}}$, leur liaison grandit et un **cluster** se forme, fusionnant la dimension symbolique et la dimension perceptive. Cette approche dispense

d'un connecteur artificiel entre un réseau neuronal et un moteur logique indépendant. C'est le **réseau lui-même**, via la mise à jour des synergies, qui assure l'unification de ces deux niveaux de représentation. On obtient ainsi une **plasticité accrue**, où les règles symboliques peuvent être ajustées ou rejetées en fonction de leur compatibilité avec les données sensorielles. Elles se reforment ou disparaissent au sein du **SCN**, s'adaptant en fonction des synergies détectées. Cette flexibilité offre un **socle pour un raisonnement abstrait ancré dans la perception**, une condition essentielle pour l'**IA Forte**, qui doit être capable d'extraire des modèles du monde tout en manipulant des structures logiques évolutives.

1.6.7.4. Défis Mathématiques : un Pas vers une Théorie de la Cognition Synergique

Le **Deep Synergy Learning (DSL)** peut se voir comme un **système dynamique** non linéaire de grande dimension, dont la **matrice** $\Omega(t)$ (ou un hyper-graphe) évolue suivant une loi similaire à

$$\Omega(t + 1) = \Omega(t) + \eta [S(\Omega(t)) - \tau \Omega(t)],$$

où $\Omega(t)$ désigne l'ensemble des pondérations $\{\omega_{i,j}(t)\}$ à l'instant t , et $S(\Omega(t))$ symbolise la mesure de **synergie** calculée entre les entités, qui peut elle-même dépendre de l'état global. Cette dynamique est susceptible de générer des **attracteurs** multiples (plusieurs organisations cognitives possibles), des **bifurcations** lorsque les synergies changent brusquement, voire des **cycles limites** interprétables comme des rythmes cognitifs ou des alternances d'états mentaux. L'**analyse** formelle (existence et stabilité d'attracteurs, transitions critiques, phénomènes d'hystérèse) reste toutefois un **défi** majeur, notamment si l'on souhaite déployer de nombreuses entités ou autoriser une définition non triviale de la synergie (information mutuelle n-aire, PID, etc.).

Pour gérer la croissance potentiellement explosive du nombre de liaisons $\{\omega_{i,j}\}$ et leur mise à jour en temps réel, on se tourne fréquemment vers des **méthodes de parcimonie** (seuils, coupes, hiérarchies de clusters) ou des **stratégies** d'échantillonnage. Certains mécanismes inspirés de la **biologie** — par exemple un terme d'oubli rapide ou une saturation — contribuent également à la **scalabilité**. Le choix précis de la fonction de synergie S , de la règle de mise à jour et de la topologie initiale influe sur la façon dont le réseau s'organise.

En parallèle, les travaux en **sciences cognitives** ou neurosciences computationnelles (comme la **théorie de l'information intégrée** de Tononi) proposent de quantifier la **part** d'information réellement unifiée dans un système. Dans l'esprit du **DSL**, on peut définir une **intégration globale** à partir de la **somme** des synergies (ou d'une co-information n-aire) :

$$\text{Intégration globale} = \sum_{i,j} \omega_{i,j} S(\mathcal{E}_i, \mathcal{E}_j),$$

ou encore effectuer une version plus générale au sein d'un **cluster** n-aire. Cette démarche veut évaluer la "**richesse**" émergente dans un réseau en perpétuelle auto-organisation. Elle ouvre un **pas** vers des réflexions sur la **conscience artificielle** ou l'**autonomie** inspirée de la cognition humaine, puisqu'elle quantifie à quel degré le réseau regroupe et **intègre** l'information sous forme de micro-réseaux cognitifs, susceptibles d'exhiber des propriétés d'**émergence** et de **complexité** caractéristiques des systèmes vivants.

1.7. Défis, Contraintes et Ouvertures

Les sections précédentes (1.1 à 1.6) ont mis en lumière la **logique fondamentale** du **Deep Synergy Learning (DSL)**, ses **mécanismes** d'auto-organisation et ses **capacités** à apporter des solutions nouvelles dans de multiples domaines applicatifs (vision, audio, robotique, recommandation, diagnostic, etc.). Toutefois, comme toute approche novatrice, le DSL n'échappe pas à des **défis** et des **contraintes** majeurs. Sur le plan **computationnel**, **théorique**, **méthodologique** et **éthique**, de nombreuses questions demeurent ouvertes, exigeant une recherche soutenue pour perfectionner le paradigme et garantir son insertion fluide dans des environnements industriels, médicaux, ou encore scientifiques.

La volonté du DSL de fonctionner sans architecture hiérarchique figée, et de laisser les entités s'**auto-organiser** via des synergies, requiert notamment :

- Une **scalabilité** adéquate (comment passer à des milliers ou millions d'entités sans explosion de la complexité ?),
- Une **qualité** et une **accessibilité** suffisantes des données (pour estimer la synergie de façon robuste),
- Des **algorithmes d'optimisation** aptes à gérer des systèmes dynamiques, non linéaires et potentiellement hétérogènes,
- Un **contrôle** de la stabilité et de la convergence (éviter les oscillations ou l'effondrement du réseau),
- Une **interprétabilité** pour l'humain, d'autant plus cruciale lorsque les décisions impactent des vies ou des processus critiques,
- Des **considérations éthiques** et réglementaires pour encadrer l'usage de systèmes auto-organisés, notamment en termes de responsabilité, d'acceptabilité, ou de gestion de biais,
- Des **comparaisons expérimentales** rigoureuses avec les approches existantes (réseaux neuronaux profonds, méthodes symboliques, algorithmes d'optimisation classiques) pour valider la pertinence et les conditions d'efficacité du DSL.

Dans cette section (1.7), nous examinons un à un ces **défis**, **contraintes** et **ouvertures**, déclinés en sept points :

- **Complexité Computationnelle et Scalabilité (1.7.1)**
- **Qualité et Disponibilité des Données (1.7.2)**
- **Développement d'Algorithmes d'Optimisation Appropriés (1.7.3)**
- **Contrôle et Stabilité des Processus Auto-Organisés (1.7.4)**
- **Interprétabilité et Explicabilité pour l'Humain (1.7.5)**
- **Considérations Éthiques et Réglementaires (1.7.6)**

- **Comparaisons Expérimentales avec d’Autres Approches (1.7.7)**

Ce panorama permettra de saisir la **complexité** inhérente à l’approche synergique, tout en mettant en relief les **axes** de recherche et les **collaborations** scientifiques qu’elle appelle, pour faire du DSL un cadre de plus en plus solide dans l’écosystème de l’IA et de l’apprentissage automatique.

1.7.1. Complexité Computationnelle et Scalabilité

La première critique ou préoccupation qu’on peut formuler à l’égard du **DSL** concerne la **taille** potentielle du **graphe** des entités et la **complexité** des opérations nécessaires pour **calculer** ou **mettre à jour** les **pondérations** synergiques ($\omega_{i,j}$). En effet, si l’on considère n entités $\{\mathcal{E}_1, \dots, \mathcal{E}_n\}$, avec une **synergie binaire** $S(\mathcal{E}_i, \mathcal{E}_j)$ pour chaque paire (i, j) et une **mise à jour régulière**, la **complexité brute** peut rapidement atteindre $O(n^2)$ par itération. Cette croissance quadratique devient impraticable lorsque n atteint plusieurs milliers ou millions de composantes. De plus, si l’on veut aller vers des **synergies n-aires** (pour capturer des coopérations à plusieurs entités simultanément), la situation s’aggrave encore (complexité exponentielle).

1.7.1.1. Problématique du “Tout Relier à Tout”

L’un des principes centraux du **Deep Synergy Learning (DSL)** est de laisser les entités $\{\mathcal{E}_i\}$ se connecter ou se déconnecter au fil du temps, suivant la **synergie** perçue. Pourtant, permettre qu’une entité puisse se relier à toutes les autres sans limite conduit vite à un **graphe complet**, où chaque paire (i, j) exige un calcul de synergie $S(\mathcal{E}_i, \mathcal{E}_j)$. Cette démarche devient rapidement ingérable :

- Le calcul de $S(\mathcal{E}_i, \mathcal{E}_j)$ pour toutes les paires (i, j) croît en $O(n^2)$.
- La mise à jour des pondérations $\omega_{i,j}$ (section 1.4.5) présente la même complexité en $O(n^2)$.
- Le stockage de l’ensemble $\{\omega_{i,j}\}$ requiert une **mémoire** en $O(n^2)$.

Dans des applications **massives** (robotique multi-capteurs, systèmes de recommandation avec des millions d’utilisateurs et de contenus, vision distribuée avec d’innombrables descripteurs), la construction d’un tel graphe complet compromet la **scalabilité** du DSL. Il devient crucial de maintenir une certaine **parsimonie** au sein du réseau, de sorte à ne pas créer aveuglément des liaisons $\omega_{i,j}$ pour chaque paire.

Une **stratégie** commune consiste à **restreindre** le calcul de synergie $S(\mathcal{E}_i, \mathcal{E}_j)$ aux seules paires dont les entités se révèlent “assez proches” ou déjà **corrélées**. On peut, par exemple, refuser la création de liens si la distance entre \mathbf{x}_i et \mathbf{x}_j dépasse un certain **seuil** ϵ . On peut aussi confiner le calcul à un **voisinage** k-plus-proches-voisins, limitant la croissance de connections. Par ailleurs, il faut un **mécanisme** pour découvrir graduellement de nouvelles liaisons, via un **échantillonnage** ou une **partition** initiale, sans explorer d’emblée toutes les paires (i, j) .

En parallèle, on définit des **règles** pour **supprimer** ou **inactiver** automatiquement les liens jugés trop faibles ou non pertinents (auto-suppression sous un seuil ω_{\min}). De cette manière, la densité du graphe reste modérée. On vise alors une complexité proportionnelle au nombre de liaisons

“actives”, ce qui peut être bien inférieur à n^2 . Au bout du compte, le **réseau** se rapproche d’un **k-NN** ou d’un ϵ -**radius** dynamique, sur lequel on applique les mises à jour (section 1.4.5). Le **DSL** y opère donc ses principes d’auto-organisation et de synergie, tout en évitant l’écueil d’un graphe complet ingérable.

1.7.1.2. Calcul de Synergie : Coût des Mesures Informationnelles

Lorsque la **synergie** prend la forme d’une **co-information** ou d’une **information mutuelle** entre entités modélisées comme variables aléatoires, on se heurte à la difficulté du calcul ou de l’estimation de grande dimension. Les estimateurs non paramétriques (k-NN, Kernel density) sont souvent coûteux et la quantité de données exigée peut augmenter de manière exponentielle avec la dimension (curse of dimensionality\text{curse of dimensionality}). Pour rendre ces **métriques d’information** plus utilisables en pratique, plusieurs stratégies sont envisageables. Il est possible de **limiter** la dimension en extrayant des **features** plus compactes à l’aide d’un autoencodeur ou d’une PCA, ce qui réduit la charge des estimateurs d’entropie et de co-information. Une autre option consiste à **recourir** à des mesures de synergie plus **légères**, comme la distance euclidienne ou la similarité cosinus, même si cela implique une perte de richesse par rapport à l’information mutuelle stricte. Enfin, lorsqu’il est essentiel de conserver une approche basée sur l’**information théorique**, on peut **employer** des heuristiques d’approximation, comme le **downsampling** des données, une modélisation paramétrique (Gaussienne, mixture Gaussienne) ou des estimateurs simplifiés.

Ces méthodes aident à concilier la **vision** d’un DSL basé sur la co-information et la **réalité** des applications de grande dimension, pour lesquelles un calcul exact d’entropie resterait prohibitif.

1.7.1.3. Mise à Jour Itérative ou Partielle

Même si l’on parvient à maintenir une **structure parcimonieuse** (où un sous-ensemble restreint de liaisons $\omega_{i,j}$ demeure actif), la **mise à jour** de chaque pondération peut néanmoins devenir un **processus lourd** dans un réseau de grande taille. Pour rendre le **Deep Synergy Learning (DSL)** praticable à grande échelle, il est donc nécessaire de recourir à des **méthodes itératives** ou à des **mises à jour partielles**, évitant un recalcul exhaustif à chaque itération.

Une solution envisageable réside dans une **mise à jour locale en parallèle**. Chaque entité \mathcal{E}_i ne s’occupe que de ses **voisins** actuels, c’est-à-dire les nœuds auxquels elle est reliée par des liaisons $\omega_{i,j}$ non négligeables. Chacune met à jour les pondérations de ses connexions en parallèle avec les autres entités. Cette approche repose sur une **communication distribuée** ou une **mutualisation des ressources** (GPU, TPU), tout en restant locale. Seuls les liens existants sont explorés, évitant ainsi un balayage exhaustif de toutes les paires (i, j) .

Il demeure toutefois essentiel d’introduire un **mécanisme** pour éviter les conflits et saturations éventuelles. Par exemple, si plusieurs entités cherchent à coopérer en excès avec la même cible, on peut concevoir une **inhibition compétitive** assurant que les liaisons se stabilisent autour d’un certain nombre de voisins. Certaines idées issues des **systèmes multi-agents** (communication locale asynchrone, règles de stabilisation) favorisent la **convergence** du réseau. Ce paradigme

s'apparente finalement à une **simulation** de réseaux biologiques ou neuronaux, où les neurones mettent à jour leurs connexions en temps réel de manière distribuée. La **scalabilité** demeure possible tant que la **densité** du graphe demeure limitée et que les calculs de synergie (ou de similarité) ne concernent pas l'ensemble des paires dans un espace de grande dimension.

1.7.1.4. Vers une Mathématique de la Parcimonie et du Grouping Évolutif

Une autre voie de recherche consiste à définir une **énergie** ou un **coût** global $\mathcal{J}(\Omega)$ pour le réseau Ω , englobant non seulement le terme $-\sum \omega_{i,j} S_{i,j}$ (valorisant les liens qui affichent une forte synergie) mais également un terme de **pénalisation** du nombre de liaisons non nulles. On peut par exemple écrire :

$$\mathcal{J}(\Omega) = - \sum_{i,j} \omega_{i,j} S_{i,j} + \alpha \|\Omega\|_0,$$

où $\|\Omega\|_0$ compte le nombre de connexions actives (i.e. $\omega_{i,j} > 0$) et $\alpha > 0$ contrôle l'importance de la parcimonie. Cette formulation **force** le réseau à **rester** relativement épars, tout en **maximisant** la somme de synergies. L'analyse mathématique des **minima locaux** de $\mathcal{J}(\Omega)$ et la description des **trajectoires** de descente dans un espace de grande dimension soulèvent des **défis majeurs**. Il s'agit de caractériser la **stabilité** des clusters émergents ainsi que la **croissance** potentielle de la taille du réseau au fil des itérations. Cette approche ouvre des perspectives pour un contrôle plus strict de la **complexité**, en assurant qu'une connectivité limitée ne se transforme pas en un graphe complet, préservant ainsi la faisabilité du **DSL** à grande échelle.

Conclusion

La **complexité computationnelle** et la **scalabilité** constituent sans doute les **premiers** grands défis pour le **Deep Synergy Learning**. Sans mesures de parcimonie et sans heuristiques de mise à jour partielle, le coût en temps et en mémoire peut devenir prohibitif. Les **pistes** de solutions incluent :

- **Sparsité** intégrée dans la création/rupture de liens,
- **Méthodes** d'estimation approximative des synergies dans des espaces de haute dimension,
- **Parallélisme** et algorithmes distribués,
- **Formulations** d'énergie globale favorisant la parcimonie.

Ces recherches sont cruciales pour permettre au DSL de s'étendre à grande échelle (milliers, millions d'entités) sans perdre l'efficacité ou la réactivité nécessaires dans les applications industrielles, médicales, ou cognitives.

1.7.2. *Qualité et Disponibilité des Données*

Au-delà des défis de **complexité** et de **scalabilité** (1.7.1), le **Deep Synergy Learning (DSL)** doit aussi relever des enjeux majeurs quant à la **qualité** et la **disponibilité** des données. En effet, la logique synergique et auto-organisée repose sur l'évaluation locale de **mesures de synergie** ($S(\mathcal{E}_i, \mathcal{E}_j)$) et sur la **mise à jour** (positive ou négative) de pondérations $\omega_{i,j}$. Lorsque les données sont trop **bruitées**, trop **parcellaires** ou manquent de **représentativité**, la convergence vers des clusters pertinents ou la stabilité de l'architecture peut être compromise. Le DSL, tout en étant plus tolérant aux lacunes qu'un réseau neuronal figé, n'en demeure pas moins dépendant d'un certain **niveau** de fiabilité et de diversité des entrées pour exploiter la synergie.

Dans cette sous-section (1.7.2), nous analyserons :

- Comment la **qualité** et la **couverture** des données impactent le calcul de la synergie,
- Pourquoi la **disponibilité** (flux continu vs. données rares) peut influencer sur la dynamique du DSL,
- Les biais possibles et les problèmes d'**incohérence** qui peuvent fausser l'auto-organisation,
- Les pistes pour **sécuriser** ou **améliorer** la robustesse de l'apprentissage synergique face à des données imparfaites.

1.7.2.1. *Importance d'une Bonne Couverture et Diversité des Données*

Le **Deep Synergy Learning (DSL)** s'appuie sur la fonction de synergie $S(\mathcal{E}_i, \mathcal{E}_j)$ pour guider la mise à jour des pondérations $\omega_{i,j}(t)$. Lorsque ces données \mathcal{E}_i et \mathcal{E}_j sont rares, bruitées ou biaisées, l'estimation de leur **distance** ou **similarité** (ou même de leur **co-information**) risque d'être inexacte. Cela peut entraîner un renforcement abusif de liens qui ne recouvrent aucune complémentarité réelle ou, à l'inverse, un affaiblissement injustifié de liaisons potentiellement utiles. Le réseau peut dès lors former des **clusters** peu pertinents, traduire des **biais** préexistants dans les données ou négliger des sous-groupes minoritaires.

Au-delà du simple problème d'**apprentissage** unitaire, le **DSL** met en relief la **couverture** temporelle. Si les conditions (saisons, configurations) varient, un système cherchant à s'adapter en continu aura besoin d'exemples suffisants pour chaque **régime** de données (périodes spécifiques, sous-populations) afin de réévaluer les synergies. Si l'on ne présente jamais de situations "hiver" lors de la constitution initiale du réseau, il se peut que les liaisons nécessaires ne s'instaurent pas, et le DSL échouera à former des **clusters** adaptés dès l'apparition d'un climat hivernal. Cette logique s'applique également à d'autres domaines. Un **flux** d'information continu exige un nombre d'**itérations** suffisant pour que chaque nouvelle modalité ou contexte s'intègre progressivement dans le **Synergistic Connection Network**, consolidant les liens synergiques pertinents.

La diversité et la **représentativité** des données apparaissent donc cruciales pour qu'un **DSL** puisse véritablement tirer parti de son **auto-organisation**. Sans cette diversité, il risque de manifester des **biais** similaires à ceux de nombre de méthodes d'apprentissage, la différence étant qu'ici c'est la

formation des clusters (et la dynamique des liaisons ω) qui pâtit des lacunes dans la couverture des phénomènes et des populations.

1.7.2.2. Bruit, Lacunes et Données Incomplètes

Comme discuté en section 1.5.3, le **Deep Synergy Learning (DSL)** fait preuve d'une **flexibilité** remarquable vis-à-vis des données incomplètes ou bruitées. Une entité trop incertaine finit par se **retrouver isolée** si sa synergie avec les autres entités demeure inexistante, préservant ainsi la **structure** principale du réseau. Dans ce cas, le réseau peine à structurer une organisation cohérente. Lorsque la majorité des entités est affectée par un **bruit excessif** ou un **manque de données**, l'**auto-organisation** ne parvient pas à dégager de **synergies réelles**. Les pondérations $\omega_{i,j}$ restent faibles, empêchant la formation de **clusters solides**, ou bien elles se renforcent à tort en raison d'**artefacts de bruit**, entraînant des structures biaisées.

Le **DSL** reste donc partiellement tributaire d'un **nettoyage** ou d'une **consolidation** préliminaire si le **bruit** se montre trop massif, sous peine de limiter sa capacité à détecter les vraies complémentarités. Plusieurs stratégies peuvent être mises en place pour améliorer la **qualité des données** et renforcer la cohérence du réseau. Un **filtrage statistique** permet d'écarter ou de sous-pondérer les entités dont les mesures sont jugées trop incertaines, en raison d'un taux élevé de valeurs manquantes ou d'incohérences répétées. Une **fusion locale** des entités fortement **corrélées** améliore la **fiabilité** du réseau : lorsque deux entités \mathcal{E}_i et \mathcal{E}_j possèdent des **données redondantes**, leur **agrégation** en une entité plus **robuste** réduit la dispersion des pondérations $\omega_{i,j}$. L'inclusion de **scores de fiabilité** dans la fonction de synergie $S(\mathcal{E}_i, \mathcal{E}_j)$ limite l'influence des estimations reposant sur un faible échantillon de points ou sur des valeurs suspectées de bruit. La similarité ou la distance peut ainsi être modulée en fonction d'un **coefficient de confiance** propre à chaque entité ou à chaque mesure.

Grâce à ces mécanismes, on endigue l'effet d'un bruit omniprésent tout en préservant la logique d'**auto-organisation** du DSL. Ainsi, même si certaines entités restent partiellement incomplètes, elles peuvent trouver leur place dans le **Synergistic Connection Network** dès lors qu'elles entretiennent au moins quelques synergies robustes avec d'autres entités.

1.7.2.3. Accès et Disponibilité Continue : le DSL comme Système en Ligne

Dans de nombreux environnements, il est nécessaire de gérer un **flux** ininterrompu de données et de réajuster en permanence les structures d'apprentissage. Le **Deep Synergy Learning (DSL)**, déjà décrit comme un processus de mise à jour itérative (voir la section 1.4.5), peut fonctionner en **mode "online"** si l'on s'assure que chaque entité \mathcal{E}_i reçoit ses observations de manière régulière ou semi-régulière. Lorsque de nouvelles mesures sont disponibles, la **synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$ est recalculée localement (ou approximée sur une fenêtre glissante), puis la pondération $\omega_{i,j}(t + 1)$ s'actualise selon l'équation :

$$\omega_{i,j}(t + 1) = \omega_{i,j}(t) + \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)].$$

Cette logique d'évolution incrémentale requiert un **accès** continu aux flux entrants, que l'on traite idéalement par un système de gestion de messages (type broker Kafka) ou par un stockage minimal assurant la conservation des $\omega_{i,j}(t)$. Le réseau se reconfigure au fil de la réception de nouvelles données, sans qu'il soit nécessaire de reprendre un apprentissage exhaustif en mode batch. Dans un contexte industriel où la disponibilité des données peut être irrégulière, il arrive parfois que certaines entités n'aient aucune mise à jour pendant une période. Il faut alors conserver leurs liaisons $\omega_{i,j}(t)$ inchangées et ne pas pénaliser excessivement l'absence de mesures, conformément aux principes d'adaptation face aux lacunes de données (section 1.5.3). Ce mécanisme d'**auto-régulation** assure une **adaptabilité continue** du **DSL** face aux changements de distribution. Lorsqu'une entité devient moins pertinente dans son **environnement**, ses **liaisons** ω diminuent progressivement, l'éloignant de son cluster d'origine. Simultanément, de **nouvelles connexions** émergent avec des entités mieux alignées sur la situation actuelle, permettant au réseau de s'**ajuster dynamiquement** sans nécessiter de réapprentissage global.

Du point de vue du **déploiement**, un tel système en ligne exige une **synchronisation** ou un **parallélisme** adapté. Chaque entité ne traite que ses connexions actives, ce qui maintient la complexité à un niveau proportionnel au nombre de liaisons conservées (voir la section 1.7.1.2 sur la parcimonie). Pour des applications massivement distribuées, il est possible de faire coexister plusieurs nœuds DSL partiels et de synchroniser leurs matrices ω par échanges de sous-graphes ou de résumés statistiques. Cette configuration favorise la **scalabilité** et la **résilience**, car si un nœud local subit une interruption ou un ralentissement, le reste du réseau n'est pas forcé de s'arrêter, et les autres entités continuent à ajuster leurs pondérations. Une telle conception rend le **DSL** apte à gérer des environnements en perpétuel changement, tout en intégrant les principes d'auto-organisation et de mise à jour continue.

1.7.2.4. Biais et Incohérences Structurelles

Dans le **Deep Synergy Learning (DSL)**, les liaisons $\omega_{i,j}$ émergent localement à travers la **synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$. Lorsque les données d'entrée sont biaisées, par exemple avec une surreprésentation de certains profils (sections 1.5.3 et 1.7.2.1), le réseau a tendance à **renforcer** les clusters correspondant aux **groupes majoritaires**, tout en marginalisant des entités issues de groupes minoritaires ou insuffisamment renseignés. Dans des domaines critiques comme le **diagnostic médical** ou la **recommandation sensible**, un tel phénomène peut avoir pour effet d'**entériner** des inégalités ou de mener à des **erreurs** d'interprétation (sous-diagnostic pour certains patients, ou éviction de contenus pertinents pour certaines catégories d'utilisateurs).

À la différence d'un **modèle supervisé** où l'on surveille une **loss** globale sur l'ensemble des données, le DSL base ses ajustements sur des **mesures locales** de synergie. Même si la structure globale Ω peut finir par présenter un arrangement biaisé, la démarche d'**auto-organisation** ne possède pas toujours de mécanismes internes de correction. Les entités minoritaires voient leurs pondérations $\omega_{i,j}$ rester faibles, du fait qu'elles ne rencontrent pas suffisamment de synergie avec la majorité. Par ailleurs, si une **incohérence** surgit ou si un artefact statistique biaisé devient prépondérant, le DSL peut stabiliser un **cluster** aberrant, en l'absence de contrôle externe.

Pour pallier cette difficulté, il est indispensable de prévoir des **mécanismes** de **validation** ou de **monitoring** global. On peut, par exemple, injecter des **entités symboliques** (sections 1.5.7 et

1.7.2.3) qui portent des règles ou des normes éthiques, limitant la croissance de clusters injustes ou repérant les configurations contradictoires. On peut également imposer un **feedback** externe, ou un signal de performance englobant des critères d'équité, veillant ainsi à empêcher la fermeture de groupes déviants. La structure ω peut alors être réévaluée à la lumière de ces **contraintes** d'équité ou de cohérence, conférant au réseau un degré de régulation qui contrebalance les dérives potentielles issues de la simple coopération locale.

1.7.2.5. Conclusion

La **qualité** et la **disponibilité** des données jouent un rôle clé dans le **Deep Synergy Learning**, comme dans toute approche d'IA. Ce modèle parvient à mieux tolérer la **partialité** en isolant les entités trop bruitées et s'adapte efficacement à un **flux continu** grâce à la mise à jour locale des pondérations. Toutefois, il reste sensible aux **biais** et nécessite une **bonne représentativité** des données pour assurer une synergie pertinente.

Un **travail préparatoire** rigoureux s'impose afin d'optimiser le fonctionnement du **Deep Synergy Learning**. Il est essentiel de **filtrer** et **pondérer** la qualité des données, d'intégrer des **scores de confiance** dans l'évaluation de la synergie, et de définir une **stratégie** de streaming ou de batch mixte pour alimenter la dynamique des pondérations ω . Il convient également de surveiller l'**évolution** des clusters afin de détecter d'éventuelles anomalies ou d'anticiper la propagation de biais au sein de l'auto-organisation.

Un **pilotage rigoureux** des données est essentiel pour exploiter pleinement le **DSL**. Un système **auto-organisé** ne peut compenser des biais majeurs de collecte ou un manque de **représentativité** sans une **ingénierie des données** adéquate. Une **gouvernance robuste** garantit une qualité suffisante des flux entrants, évitant ainsi que des **liaisons synergiques** se forment sur des bases erronées ou que des **clusters** se structurent autour de signaux bruités.

1.7.3. Développement d'Algorithmes d'Optimisation Appropriés

Une fois admise la nécessité de gérer la **complexité** (1.7.1) et de soigner la **qualité** des données (1.7.2), se pose la question cruciale des **algorithmes** qui permettront de **piloter** la dynamique du **Deep Synergy Learning (DSL)**. En effet, ce paradigme repose sur l'**auto-organisation** de multiples entités $\{\mathcal{E}_i\}$ et sur la **mise à jour** de l'ensemble des pondérations synergiques $\omega_{i,j}(t)$. Dans le modèle théorique simple, on écrit souvent :

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta[S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)],$$

où $S(\mathcal{E}_i, \mathcal{E}_j)$ désigne la **synergie** entre ces entités.

En pratique, mettre en œuvre cette loi pour un grand nombre d'entités, avec des synergies potentiellement complexes (non linéaires, conditionnelles, n-aires), n'est pas trivial. On a besoin de **méthodes d'optimisation** capables de **stabiliser** le système en évitant les oscillations ou l'explosion des pondérations. Elles doivent aussi **accélérer** la convergence vers des configurations pertinentes comme les clusters et macro-clusters. L'intégration de **contraintes** est essentielle pour garantir la cohérence, la parcimonie et la compatibilité avec des règles symboliques. Enfin, ces

méthodes doivent pouvoir **s'adapter** en temps réel à l'apparition de nouvelles entités ou de nouveaux flux de données.

Cette sous-section (1.7.3) discute quelques **pistes** mathématiques et algorithmiques pour rendre le DSL plus solide et plus efficace :

- Des méthodes **locales** (mise à jour distribuée),
- Des formulations **globales** (fonction d'énergie, descente de gradient généralisée, approches de type recuit simulé),
- Des **heuristiques** ou des règles inspirées de la biologie et des systèmes complexes (sélection, reproduction, extinction),
- Des **extensions** pour la synergie n-aire ou conditionnelle.

1.7.3.1. Approche Locale et Distribué

Dans l'esprit **biologique** et de nombreux modèles de **systèmes complexes**, on peut envisager la mise à jour de chaque pondération $\omega_{i,j}$ de manière **locale**, sans qu'un contrôle global impose la synchronisation. Chaque entité \mathcal{E}_i communique uniquement avec ses **voisins** au sens du graphe, autrement dit les entités \mathcal{E}_j pour lesquelles $\omega_{i,j}(t)$ est significatif. L'entité \mathcal{E}_i calcule alors la **synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$ sur la base de mesures ponctuelles (ou d'un historique local) et applique la règle :

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta_{i,j}(t) [S_{i,j}(t) - \tau_{i,j}(t) \omega_{i,j}(t)].$$

Cette équation reprend le formalisme général du **Deep Synergy Learning** (section 1.4.5) en autorisant des coefficients $\eta_{i,j}(t)$ et $\tau_{i,j}(t)$ qui dépendent de la paire (i, j) ou varient dans le temps, procurant une **régulation** fine. L'**avantage** est de rendre l'algorithme **parallélisable**, puisque chaque entité \mathcal{E}_i peut mettre à jour ses liaisons sans solliciter un "master" central. Une difficulté réside dans la synchronisation des mises à jour. Si plusieurs entités modifient en même temps leurs liaisons synergiques, des boucles instables ou des conflits peuvent émerger. Une gestion asynchrone ou régulée des ajustements est alors nécessaire pour éviter des oscillations non maîtrisées dans le réseau. Des stratégies d'**asynchronisme** maîtrisé ou de communication par itérations (steps) successives sont alors recommandées.

Dans cette optique, on peut se référer à diverses **métaphores** ou **règles** d'inspiration biologique. Par exemple, une **mise à jour** "hebbienne" renforce la pondération $\omega_{i,j}$ lorsqu'on observe une co-activation récurrente des entités \mathcal{E}_i et \mathcal{E}_j , et l'affaiblit en l'absence de co-occurrence. Par ailleurs, on peut sélectionner ou reproduire certaines liaisons. Si des paires (i, j) restent faiblement utiles, elles sont éliminées, tandis que d'autres liens se renforcent ou se dupliquent sur d'autres nœuds. Une compétition locale peut limiter la densité des connexions en contraignant la somme $\sum_j \omega_{i,j}$ pour chaque entité \mathcal{E}_i ou en appliquant un softmax local, favorisant quelques connexions fortes plutôt qu'un réseau dilué.

Par ces variantes, on parvient à **contrôler** la **densité** du graphe, la vitesse de renforcement des pondérations et la **résilience** face au bruit ou aux données incomplètes (sections 1.7.1.2 et 1.5.3). Le **DSL** conserve sa **scalabilité** tout en respectant l'**auto-organisation**. Chaque mécanisme reste local, assurant une dynamique distribuée compatible avec des implémentations parallèles sur GPU ou TPU, ainsi qu'avec des approches multi-agents où chaque entité ajuste ses connexions en fonction de ses propres observations et interactions.

1.7.3.2. Formulation Globale via Fonction d'Énergie

Dans une perspective plus proche de l'**optimisation** (au sens des approches traditionnelles en IA ou en physique statistique), il est concevable de décrire l'**auto-organisation** du Deep Synergy Learning (DSL) par une **fonction d'énergie** $\mathcal{J}(\Omega)$. On peut écrire une expression qui généralise la somme des synergies et la pondère par un terme de **régularisation**, selon :

$$\mathcal{J}(\Omega) = - \sum_{i,j} \omega_{i,j} S_{i,j} + R(\Omega).$$

Dans ce formalisme, la somme $\sum_{i,j} \omega_{i,j} S_{i,j}$ valorise la **coopération** (au sens de la synergie, voir sections 1.4.4 et 1.4.5), tandis que le terme $R(\Omega)$ constitue une **régularisation**, par exemple $\alpha \|\Omega\|^2$ ou $\alpha \|\Omega\|_0$, qui maintient la parcimonie ou la cohérence structurelle du **Synergistic Connection Network**. Cette formulation intègre les principes abordés en 1.7.1.4, notamment le **contrôle de la densité** et le **maintien des contraintes**, au sein d'une unique fonction de coût. L'objectif global revient à **maximiser** la somme des synergies ou, en termes d'énergie, à **minimiser** \mathcal{J} .

Pour résoudre l'évolution de Ω dans ce cadre, on peut recourir à une **descente de gradient**. Si l'on note $\omega_{i,j}$ un des paramètres, sa dérivée partielle vis-à-vis de \mathcal{J} prend la forme

$$\frac{\partial \mathcal{J}}{\partial \omega_{i,j}} = - S_{i,j} + \frac{\partial R(\Omega)}{\partial \omega_{i,j}}.$$

En discrétisant cette descente, on retrouve l'équation qui met à jour $\omega_{i,j}$ de façon incrémentale :

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta [S_{i,j} - \tau \omega_{i,j}(t)],$$

ce qui explicite la proximité entre la **logique locale** décrite en 1.4.5 et la **formulation** d'une énergie globale $\mathcal{J}(\Omega)$. Dans les problèmes complexes, la descente de gradient peut aboutir à des minima locaux sous-optimaux. On peut alors envisager un **recuit simulé** (simulated annealing) directement sur l'espace du graphe Ω . Cette méthode consiste à perturber par moments la configuration courante (création ou suppression de liens), et à n'accepter ou refuser ces modifications qu'avec une probabilité liée à la variation $\Delta \mathcal{J}$ et à un paramètre de température qui décroît au fil des itérations. Dans le contexte du **DSL**, où la synergie peut être **n-aire** (voir section 1.4.7), une exploration plus large de l'espace combinatoire devient essentielle. Se limiter à des ajustements locaux risquerait d'aboutir à un **cluster** partiel, sans permettre l'émergence de synergies plus globales. Le recuit simulé ou des **algorithmes évolutionnaires** jouent alors le rôle de recherche d'une configuration plus robuste, accompagnant l'auto-organisation locale. Cette combinaison de l'approche **par**

énergie et de la **dynamique en continu** ouvre une voie pour maîtriser mathématiquement la convergence et la structure du **Synergistic Connection Network**, en particulier lorsqu'on cherche à maintenir une **parsimonie** explicite ou à inclure des contraintes topologiques (voir 1.7.1.4).

1.7.3.3. Synergie n-aire et Extensions Conditionnelles

Le **Deep Synergy Learning (DSL)** se focalise souvent sur la **coopération** entre deux entités \mathcal{E}_i et \mathcal{E}_j . Toutefois, certains problèmes exigent une **synergie n-aire**, c'est-à-dire impliquant un **groupe** $\{\mathcal{E}_{i_1}, \dots, \mathcal{E}_{i_n}\}$ de cardinal $n \geq 3$. Dans ce cas, on peut introduire des **hyper-arêtes** ω_{i_1, \dots, i_n} plutôt que de simples liaisons binaires. La mise à jour de la pondération hyper-arête s'effectue alors selon une généralisation de la loi (voir section 1.4.5), par exemple :

$$\omega_{i_1, \dots, i_n}(t+1) = \omega_{i_1, \dots, i_n}(t) + \eta [S_n(\mathcal{E}_{i_1}, \dots, \mathcal{E}_{i_n}) - \tau \omega_{i_1, \dots, i_n}(t)].$$

Dans cette équation, la fonction $S_n(\mathcal{E}_{i_1}, \dots, \mathcal{E}_{i_n})$ évalue la **synergie** globale de l'ensemble $\{\mathcal{E}_{i_k}\}$. Sa définition peut s'appuyer sur une co-information n-aire, sur un gain de performance collectif ou sur toute métrique reflétant l'apport mutuel de ce sous-groupe. Néanmoins, la **complexité** combinatoire s'en trouve rapidement démultipliée, car on doit gérer des hyper-arêtes potentiellement innombrables lorsque n augmente. Pour cette raison, il est fréquent de recourir à des **heuristiques** afin de détecter des triplets, quadruplets, etc., seulement lorsque l'on constate une coopération stable entre ces entités ; on se contente de créer l'hyper-arête ω_{i_1, \dots, i_n} après avoir observé à maintes reprises que ces n entités coopèrent ensemble.

Par ailleurs, il arrive que la **synergie** dépende d'un **contexte** \mathbf{c} . Un exemple typique survient lorsque l'association audio–texte n'est pertinente que dans une situation donnée (un lieu, une heure, une ambiance sonore). La règle de mise à jour se réécrit alors :

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta [S(\mathcal{E}_i, \mathcal{E}_j \mid \mathbf{c}(t)) - \tau \omega_{i,j}(t)].$$

Dans ce cadre, il est nécessaire de calculer un $S(\cdot \mid \mathbf{c})$ qui peut varier considérablement en fonction du contexte. La tâche d'**estimation** et d'**optimisation** devient plus complexe, car la distribution induite par $\mathbf{c}(t)$ peut évoluer au fil du temps ou présenter plusieurs régimes de fonctionnement distincts. Les algorithmes d'auto-organisation du DSL doivent alors incorporer des modèles de synergie **conditionnelle**, qui modulent la pondération $\omega_{i,j}$ suivant la valeur prise par \mathbf{c} . Cette extension demeure essentielle dans de multiples applications, comme la robotique adaptative (un même couple de capteurs se révèle pertinent ou non selon le mode de locomotion) ou le traitement multimodal (un segment audio et un mot ne coïncident que dans un contexte temporel précis). Elle requiert cependant des stratégies de **parcimonie** ou de **sampling** pour éviter une explosion combinatoire lorsqu'on veut évaluer la synergie conditionnelle sur tous les contextes possibles.

1.7.3.4. Contrôle de la Stabilité et Convergence Globale

Dans tout **système dynamique** où les éléments sont fortement couplés, des **cycles** ou des **oscillations** peuvent apparaître. Un sous-réseau peut se renforcer temporairement, avant qu'un autre ne prenne le relais en affaiblissant les liens précédents, entraînant une alternance continue. Il

peut aussi survenir un phénomène de **basculement** entre plusieurs attracteurs stables (configurations rivales) si la structure du **Synergistic Connection Network** autorise plusieurs manières de s'organiser pour la même situation. Le **Deep Synergy Learning (DSL)** n'échappe pas à ces scénarios, en particulier lorsque la fonction de synergie $S(\mathcal{E}_i, \mathcal{E}_j)$ est très "non linéaire" ou quand la synergie n-aire entre plusieurs entités génère des interactions complexes (sections 1.4.7 et 1.7.3.3).

Pour limiter ce risque d'oscillations ou de dérive, il convient d'inclure des **mécanismes** ou des **règles** de contrôle dans la mise à jour des pondérations $\omega_{i,j}$. On peut imposer un **amortissement** progressif (faire varier η ou τ dans le temps), intégrer une **pénalisation** plus agressive des liens excessivement grands, ou encore définir des **seuils de saturation** empêchant une pondération $\omega_{i,j}$ de croître sans borne. Des analyses mathématiques, telles que l'étude de la **Jacobienne** locale ou des **fonctions de Lyapunov** globales, peuvent démontrer qu'un schéma de mise à jour particulier garantit la convergence vers un état stable (ou au moins vers un cycle stable), sans risquer un chaos indésirable.

Il est également envisageable de greffer un **module** de contrôle hiérarchique (de plus haut niveau), qui surveille la dynamique globale du réseau et injecte ponctuellement des modifications de paramètres η ou τ si la variance du système dépasse un certain seuil. Ce principe d'**hybridation** (auto-organisation libre + régulation macro) rappelle certaines structures biologiques ou cognitives, où les ajustements locaux cohabitent avec des boucles de rétroaction plus centrales. On cherche alors à préserver la **liberté synergique** au niveau local, tout en s'assurant que le **réseau** ne sombre pas dans une instabilité excessive. Cet équilibre soulève des questions d'**optimisation** et de **pilotage** du **DSL**, qui restent ouvertes à la recherche, notamment sur les plans pratique (implémentation à grande échelle) et théorique (analyse de la dynamique dans un espace de forte dimension).

Conclusion

Le **développement d'algorithmes d'optimisation** pour le **Deep Synergy Learning** s'annonce particulièrement riche et complexe. Les **règles de mise à jour** naïves suffisent à illustrer le concept, mais ne passent pas toujours à l'échelle ni ne gèrent la synergie n-aire, les contraintes symboliques, ou la minimisation globale d'une fonction d'énergie. Les approches possibles incluent :

- **Mise à jour locale** inspirée de la biologie (sélection, Hebb, normalisation compétitive),
- **Descente d'énergie** globale (fonctions \mathcal{J} pénalisant la densité, favorisant la synergie),
- **Algorithmes évolutifs** (recuit simulé, heuristiques stochastiques) pour éviter les minima locaux,
- **Extensions** aux hyper-arêtes (synergies n-aires), aux synergies conditionnelles, ou aux règles symboliques,
- **Contrôles** de stabilité et d'amortissement, voire un superviseur partiel pour limiter les oscillations et injecter des priorités.

Trouver la meilleure **formulation** et le meilleur **algorithme** dépendra de l'application (vision, robotique, recommandation, etc.), du **volume** de données et de la **structure** (binaire ou n-aire, symbolique, multi-modal). La poursuite de ces recherches est incontournable pour que le **DSL** gagne en robustesse, en efficacité, et en adoptabilité dans le monde réel.

1.7.4. Contrôle et Stabilité des Processus Auto-Organisés

L'**auto-organisation** est au cœur du **Deep Synergy Learning (DSL)**, permettant à un grand nombre d'entités (capteurs, modules de calcul, représentations symboliques ou sub-symboliques) de s'agencer spontanément en **clusters** ou **macro-clusters** selon leurs synergies. Cette dynamique, bien qu'elle confère au **DSL** sa **plasticité** et sa **capacité d'adaptation**, peut également poser des **problèmes de stabilité**. Des oscillations, des attracteurs multiples ou des comportements chaotiques peuvent émerger, compliquant la convergence et l'organisation cohérente du réseau. De plus, dans bien des cas (robotique, diagnostic, logistique), il faut **contrôler** partiellement la configuration auto-organisée, ne serait-ce que pour garantir la **sécurité**, la **fiabilité**, ou la **cohérence** avec des règles externes.

Cette sous-section examine les **défis** liés à la **dynamique** d'un système **DSL** ainsi que les **mécanismes** permettant d'en assurer la **stabilité** et le **contrôle**. Un système couplé peut engendrer des **oscillations** ou des **cycles**, tandis que les pondérations synergiques ω peuvent soit **exploser**, soit **s'effondrer** sous certaines conditions. L'existence de **multiples attracteurs** en concurrence peut également mener à des configurations divergentes. Pour remédier à ces instabilités, il est possible d'introduire des **dispositifs** de rétroaction ou une **surveillance globale** afin d'assurer un pilotage **hiérarchique** léger. Enfin, il est crucial d'examiner les **conditions mathématiques** garantissant (ou non) la convergence de la dynamique du réseau.

1.7.4.1. Risques d'Oscillations et de Comportements Chaotiques

Dans le **Deep Synergy Learning**, chaque liaison $\omega_{i,j}$ obéit à une évolution itérative inspirée des principes décrits en section 1.4.5, ce qui donne lieu à un **système dynamique** d'environ $O(n^2)$ degrés de liberté dans le cas binaire. Pour comprendre comment apparaissent des **cycles** ou des **régimes oscillatoires**, il est éclairant d'étudier la **linéarisation** locale autour d'un point fixe ω^* . Supposons l'existence d'un équilibre ω^* satisfaisant un état stationnaire ; si l'on introduit une petite perturbation $\delta\omega$, la dynamique globale s'écrit de manière approchée :

$$\delta\omega(t+1) \approx J(\omega^*) \delta\omega(t),$$

où $J(\omega^*)$ désigne la **matrice jacobienne** évaluée en ω^* . En temps discret, un comportement oscillatoire ou pseudo-chaotique survient lorsque certaines **valeurs propres** de J possèdent une magnitude supérieure à 1. Cela signifie que la perturbation $\delta\omega$ se voit **amplifiée** au fil des itérations, au lieu d'être dissipée. En temps continu, la condition équivalente consisterait à observer des **parties réelles** positives dans les valeurs propres, menant à une instabilité locale ou à des boucles auto-entretenues.

Ce type de phénomène peut se traduire par des **cycles** permanents, ou même des trajectoires quasi chaotiques, contrariant la **convergence** du réseau vers un état stable. De tels régimes, bien que mathématiquement fascinants, s'avèrent peu souhaitables si l'on souhaite une structure DSL stable pour la robotique, la maintenance ou la prise de décision. C'est pourquoi on introduit souvent des **mécanismes** de stabilisation. Une première solution consiste à contrôler la “vitesse” de mise à jour en imposant une contrainte, par exemple $\eta \tau < 1$, dans le cas linéaire simplifié. On peut ensuite recourir à des **termes** non linéaires de saturation, destinés à éviter la croissance illimitée de certaines pondérations. Un exemple simple, dans un modèle continu, consiste à enrichir l'équation d'évolution par un terme $-\beta (\omega_{i,j})^3$:

$$\frac{d}{dt} \omega_{i,j} = \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}] - \beta (\omega_{i,j})^3,$$

où $\beta > 0$ modère la croissance des liaisons fortes. De manière analogue, on peut introduire un **facteur** de compétition locale (type softmax), en imposant $\sum_j \omega_{i,j} \leq K$ pour chaque entité \mathcal{E}_i , de sorte qu'une entité ne puisse pas développer des liens solides avec trop de partenaires en même temps.

Ces mesures limitent la **liberté** d'auto-organisation, mais s'avèrent cruciales pour écarter des régimes oscillatoires ou chaotiques susceptibles d'émerger dans un réseau fortement couplé et non linéaire. Elles assurent ainsi une forme de **stabilité** globale tout en préservant la possibilité de réorganisations locales dictées par la synergie. Un savant réglage des paramètres (η , τ , β , ou le budget $\sum_j \omega_{i,j} \leq K$) permet donc de maintenir le **DSL** dans une zone où le réseau demeure robuste et apte à converger, évitant des fluctuations intempestives tout en conservant l'essence auto-adaptative voulue.

1.7.4.2. Multiples Attracteurs, Convergence Incertaine

Il est fréquent qu'un **Deep Synergy Learning (DSL)** cherche à maximiser, ou à minimiser l'opposé, d'une **somme de synergies** $\sum_{i,j} \omega_{i,j} S_{i,j}$ éventuellement modulée par un terme de pénalisation ou de parcimonie (voir section 1.7.1.4). Dans un espace de grande dimension, on peut rencontrer de **multiples attracteurs** correspondant à des configurations dont l'énergie (ou le coût) est quasi équivalente. Ce phénomène se manifeste lorsqu'il existe plusieurs manières d'**arranger** les entités en clusters, chacune procurant un gain global proche. Le réseau peut alors converger vers l'un ou l'autre de ces attracteurs selon les conditions initiales ou de légères fluctuations.

Dans la pratique, il arrive que des “sous-réseaux” rivaux émergent au cours de la dynamique, chacun rassemblant certaines entités en un **cluster** distinct. Une compétition peut alors s'installer entre ces sous-réseaux, induisant des **instabilités** transitoires et des oscillations (voir section 1.7.4.1). Si l'on désire un **état** unique et stable, par exemple pour un robot qui ne peut se scinder en deux stratégies opposées, on peut injecter un **signal global** orientant le réseau vers l'attracteur souhaité. Cette intervention se formalise en modifiant la fonction de synergie ou en ajoutant un **feedback** de récompense plus élevé pour une configuration A que pour une configuration B. On obtient ainsi un **pilotage** macro qui restreint la pure auto-organisation, mais assure qu'en présence de minima multiples, la configuration la plus souhaitable selon les critères d'usage devienne

prépondérante. Cette démarche montre le besoin de compromis entre la liberté synergique propre au DSL et la nécessité d'un **contrôle** hiérarchique dans certaines applications critiques.

1.7.4.3. Approches Hiérarchiques Légères pour Guider l'Auto-Organisation

Dans des environnements critiques, tels que le domaine **médical** ou les **transports**, il n'est pas toujours acceptable de laisser un réseau en **Deep Synergy Learning (DSL)** se reconfigurer librement sans un regard d'ensemble. Afin de maîtriser la sécurité et la fiabilité, on introduit souvent une **couche** ou un **module** de supervision qui contrôle globalement la structure Ω . Ce superviseur observe la **matrice** des liens actifs et vérifie des **contraintes** imposées par l'application. Par exemple, on peut exiger que la somme $\sum_j \omega_{i,j}$ attachée à une entité \mathcal{E}_i ne dépasse pas un seuil spécifique, ou que certaines entités déclarées incompatibles (règles de sécurité, normes médicales contradictoires) ne figurent pas dans le même cluster. Si ce superviseur détecte une violation, il peut forcer la **mise à zéro** de certaines liaisons $\omega_{i,j}$, neutralisant ainsi des configurations illégitimes.

On peut aussi concevoir un **schéma** où la **couche basse** demeure l'auto-organisation locale et distribuée, tandis qu'une **couche haute** (un planificateur ou un décideur global) injecte à intervalles réguliers des **influx** spécifiques dans le réseau. Ces influx peuvent prendre la forme d'une modification ponctuelle de $\omega_{i,j}$, d'un renforcement sélectif d'une synergie, ou de l'assignation de priorités à certaines entités. Cet apport d'information agit comme une **guidance** de haut niveau, combinée à la dynamique adaptative du DSL.

D'un point de vue mathématique, on peut l'interpréter comme l'ajout d'un **terme** exogène dans l'équation de mise à jour, par exemple $\Delta\omega_{i,j}^{(\text{macro})}$ reflétant la volonté du **planificateur**. Cela permet d'enrichir la structure Ω sans remettre en cause l'intégralité du principe d'**auto-organisation** locale. On obtient un **compromis** dans lequel la dynamique $\omega_{i,j}$ se déroule principalement selon la synergie perçue, tout en restant arrimée à des objectifs de **performance** ou de **sécurité** plus vastes. Dans ce type d'architecture dite « hybride », le contrôle hiérarchique léger se borne à préserver la cohérence, tandis que la plupart des adaptations opérationnelles sont gérées par la mise à jour distribuée des liaisons.

1.7.4.4. Outils Mathématiques et Théoriques pour la Stabilité

Il est souvent utile de modéliser l'**évolution** d'un **Deep Synergy Learning (DSL)** sous la forme d'un **système dynamique** $\omega(t+1) = \omega(t) + F(\omega(t))$. Dans cette notation, $\omega(t)$ regroupe l'ensemble des pondérations $\omega_{i,j}(t)$, et $F(\omega(t))$ décrit la contribution que la synergie $S(\mathcal{E}_i, \mathcal{E}_j)$ (ainsi que les termes de décroissance, de régularisation, etc.) apporte à la variation entre t et $t+1$.

Pour **étudier** la stabilité locale autour d'un **point fixe** ω^* (ou d'une configuration fixe du réseau), on réalise classiquement une **linéarisation** de F . On calcule la **matrice jacobienne** :

$$J(\omega^*) = \frac{\partial F}{\partial \omega}(\omega^*).$$

Si l'on note $\delta\omega(t) = \omega(t) - \omega^*$, alors pour de petites perturbations, on obtient un modèle linéarisé :

$$\delta\omega(t+1) \approx \delta\omega(t) + J(\omega^*) \delta\omega(t).$$

En régime de temps discret, la **stabilité** locale autour de ω^* se produit si toutes les **valeurs propres** de $I + J(\omega^*)$ ont une magnitude strictement inférieure à 1. Dans un cadre continu (où l'on écrit $d/dt \omega = F(\omega)$), il faut que toutes les parties réelles des valeurs propres de $J(\omega^*)$ soient **négatives**. Cette condition s'interprète physiquement. De petites déviations $\delta\omega$ s'amortissent progressivement au fil du temps, garantissant ainsi une stabilité locale du système. Elle dépend du **choix** de la synergie $S(\mathcal{E}_i, \mathcal{E}_j)$, de la taille du pas η , des paramètres de décroissance τ , ou des non-linéarités supplémentaires (termes de saturation, etc.).

Parfois, on peut exhiber une **fonction de Lyapunov** $\Phi(\omega)$ qui décroît strictement à chaque mise à jour, prouvant la **convergence** vers un **unique attracteur**. Cela exige que la définition de la synergie ne comporte pas de **non-linéarités** trop complexes, ou que la structure ne s'étende pas à des hyper-arêtes de dimension élevée (voir la section 1.4.7 sur la synergie n-aire). Dans des **configurations** plus simples (synergie linéarisable, réseau symétrique, etc.), on peut démontrer rigoureusement l'existence et l'unicité d'un état stationnaire stable. En revanche, pour des **formulations** plus riches, où la synergie est hautement non linéaire, dépendante du contexte ou définie sur des hyper-arêtes n-aires, établir une **preuve** de stabilité globale devient plus complexe. La dynamique peut alors exhiber plusieurs attracteurs ou adopter des régimes oscillants.

Ces outils (linéarisation jacobienne, fonctions de Lyapunov) offrent néanmoins un **cadre** théorique solide pour analyser la mise à jour des pondérations $\omega_{i,j}$ dans un **DSL**, en particulier lorsque l'on cherche à garantir un comportement prédictible et éviter les **oscillations** chaotiques ou l'explosion des liaisons. Ils offrent également des **repères** pour la **conception** du système. En ajustant les paramètres η et τ ou en imposant des contraintes sur la somme et la norme des pondérations, on peut maintenir la trajectoire de ω dans une **région stable**. Cela garantit une **convergence réelle** du système, évitant ainsi toute dérive ou oscillation entre plusieurs configurations.

Conclusion

Assurer le **contrôle** et la **stabilité** des **processus auto-organisés** dans le **DSL** est un enjeu fondamental pour prévenir l'apparition de **comportements oscillants** ou chaotiques. La gestion de **multiples attracteurs** permet d'orienter l'état final et d'éviter des trajectoires divergentes. Dans certains cas, un **encadrement** minimal peut être nécessaire, sous la forme d'un module de supervision ou de règles hiérarchiques.

Les solutions mathématiques ou algorithmiques incluent la **conception** de règles de mise à jour intégrant amortissement, saturation ou normalisation pour limiter la croissance excessive des pondérations. L'introduction d'une **énergie globale** $\mathcal{J}(\Omega)$, optimisée par descente ou recuit simulé, permet d'encadrer l'évolution du système.

Un **pilotage macro** peut être mis en place en ajustant η et τ , en injectant des priorités ou en filtrant certaines liaisons. Enfin, l'**analyse de stabilité** via Jacobienne locale ou Lyapunov peut garantir, au moins localement, la convergence et éviter l'explosion des pondérations.

Maîtriser l'**auto-organisation** ne signifie pas brider la flexibilité du DSL, mais offrir à l'architecture la **robustesse** nécessaire pour opérer en conditions réelles, où l'on ne peut se permettre un comportement erratique ni une multiplicité de configurations irréconciliables.

1.7.5. Interprétabilité et Explicabilité pour l'Humain

Un objectif crucial dans la plupart des applications d'**intelligence artificielle** est de fournir non seulement des **résultats** (prédictions, recommandations, décisions), mais également une **compréhension** de la logique ou du **processus** qui y aboutit. Les méthodes classiques de **Deep Learning** (CNN, RNN, Transformers) sont souvent décriées pour leur caractère "**boîte noire**", dans la mesure où l'information se diffuse à travers des poids distribués dans de multiples couches, rendant l'**explication** pour l'humain délicate.

Le **Deep Synergy Learning (DSL)**, par son **architecture** en **entités** et en **liaisons synergiques** ($\omega_{i,j}$), offre un cadre potentiellement plus **transparent**. Il permet d'identifier quelles **entités** interagissent fortement, comment se forment les **clusters** ou **macro-clusters**, et pourquoi ces regroupements émergent sur la base d'une **synergie mesurable**. Cependant, assurer une **interprétabilité complète** n'est pas garanti. Il est nécessaire de développer des **méthodes** permettant d'extraire des **explications compréhensibles**. Cette sous-section (1.7.5) examine les **défis** et **solutions** envisageables pour améliorer l'**explicabilité** du DSL.

1.7.5.1. D'un Modèle Hiérarchique Opaque à un Réseau de Liens Synergiques

Il est courant que les **réseaux neuronaux profonds** (CNN, RNN, Transformers) diluent leur représentation interne à travers plusieurs **couches** qui transforment graduellement les descripteurs, rendant ainsi les décisions finales difficiles à expliquer. Les neurones internes ne possèdent pas d'**identité propre**. Ils agissent comme de simples réceptacles de **poids** et d'**activations**, sans que l'on puisse leur attribuer un **sens direct** ou une **fonction explicite**. À l'inverse, le **Deep Synergy Learning (DSL)** introduit un **réseau** où chaque **nœud** \mathcal{E}_i possède une **identité** clairement définie (il peut s'agir d'un **capteur** particulier, d'un **concept** sémantique, d'un **patch** visuel, d'une **règle** symbolique, etc.), tandis que les **liaisons** $\omega_{i,j}$ se **reconfigurent** progressivement, au gré de la synergie détectée entre \mathcal{E}_i et \mathcal{E}_j .

Cette structure présente plusieurs **avantages**. Elle améliore la **lisibilité** des **sous-ensembles** en mettant en évidence des **clusters** (chapitre 1.4.3) formés par des entités aux liaisons fortes, révélant une **cohérence** sous-jacente. Un cluster peut, par exemple, regrouper des entités "visuel + audio" en reconnaissance multimodale ou "utilisateur + contenus + tag de genre" dans un système de recommandation. Chaque **entité** \mathcal{E}_i possède une **signification explicite** (patch d'image, concept symbolique), facilitant ainsi l'**interprétation** des décisions du réseau. Un **macro-cluster** peut traduire un regroupement plus large, tel qu'un ensemble "groupe d'utilisateurs amateurs de jazz + contenus jazz + soirées concert" sur une plateforme culturelle, ce qui permet d'expliquer une recommandation de manière transparente : "puisque vous faites partie du cluster jazz, ce concert vous est suggéré".

Cependant, cette **lisibilité** accrue sur le plan conceptuel s'accompagne aussi d'une **complexité** potentiellement élevée. Le réseau DSL peut afficher un grand nombre de **nœuds** (surtout en contexte industriel ou multi-sensoriel), et chaque nœud peut avoir plusieurs **connexions** significatives, aboutissant à une **topologie** complexe, souvent composée de **clusters** imbriqués ou de multiples liens réciproques. Une simple **visualisation** naïve des poids $\omega_{i,j}$ peut alors ne rien révéler de clair, s'il n'existe pas de mécanisme de filtrage ou de simplification. Il devient nécessaire d'utiliser des **outils** de **hiérarchisation**, de **regroupement** ou de **filtrage** pour dériver des **chemins** explicatifs, par exemple en extrayant les principales **chaînes** de liaison qui influencent la décision finale. C'est là que réside l'enjeu d'une **explicabilité** réellement transparente dans un réseau DSL riche, qui exige de se doter de techniques aptes à condenser le graphe et à exposer les assemblées cruciales, tout en évinçant la masse de connexions moins pertinentes.

1.7.5.2. Extractions de Chemins Synergiques et Clusters Pivot

Dans un **réseau** construit selon les principes du **Deep Synergy Learning (DSL)**, il est possible d'expliquer une décision ou une recommandation en identifiant un **chemin** significatif reliant deux entités \mathcal{E}_a et \mathcal{E}_b . On peut ainsi considérer une suite :

$$\mathcal{E}_a \rightarrow \mathcal{E}_x \rightarrow \mathcal{E}_y \rightarrow \mathcal{E}_b,$$

au long de laquelle la **synergie** s'avère forte. On peut évaluer cette force par un produit du type $\omega_{a,x} \omega_{x,y} \omega_{y,b}$, exprimant la continuité du lien depuis \mathcal{E}_a jusqu'à \mathcal{E}_b via \mathcal{E}_x et \mathcal{E}_y . Par exemple, si \mathcal{E}_a représente un **utilisateur** et \mathcal{E}_b un **contenu** recommandé, un **chemin** pertinent peut inclure un **tag** musical \mathcal{E}_x et un **artiste** \mathcal{E}_y . La synergie met alors en évidence que l'utilisateur \mathcal{E}_a est historiquement associé à ce tag rock. Ce tag entretient une forte connexion avec l'artiste \mathcal{E}_y , qui est directement relié au **contenu** \mathcal{E}_b . L'explication devient un **raisonnement local**. L'utilisateur est associé au tag rock, ce tag est fortement lié à l'artiste \mathcal{E}_y , et cet artiste mène au contenu \mathcal{E}_b proposé.

Dans la pratique, on limite souvent la longueur du **chemin** afin de conserver une **explication** concise et compréhensible, tout en montrant comment la synergie chemine dans le **Synergistic Connection Network**.

Une autre forme d'explication repose sur le **cluster** (ou **macro-cluster**) auquel appartient l'entité \mathcal{E}_b . On met en avant les **entités dominantes** de ce cluster. En diagnostic médical, cela peut correspondre à un ensemble d'**indicateurs** comme la fièvre, une anomalie sanguine et une image radio suspecte, qui forment un sous-réseau fortement interconnecté. L'**interprétation** découle alors de la **cohérence interne** du **cluster**. Ces signes, fortement liés par la synergie, renforcent la conclusion de pathologie. Cette méthode est très usitée lorsque la configuration ne s'explique pas par un unique chemin, mais par un **ensemble** d'interactions locales qui se cristallisent en un groupement solidaire. L'analyse du cluster met en avant les entités centrales et leur degré de liaison ω , rendant la décision plus **transparente**.

1.7.5.3. Couplage Symbolique pour des Justifications de Haut Niveau

Dans la section 1.5.7, il a été expliqué comment intégrer des **règles** ou des **concepts** symboliques au sein du **Deep Synergy Learning (DSL)** de manière à fusionner des entités strictement sub-symboliques (features perceptuelles, segments, etc.) et des entités logico-sémantiques. Cette intégration prend toute son importance pour l'**explicabilité**, car elle autorise la production de justifications directement lisibles, sans nécessiter un métamodèle séparé. Si l'on imagine une **règle** $\mathcal{E}_{\text{rule}}$ affirmant "cercle + chiffre \Rightarrow panneau de limitation", la présence d'un lien fort $\omega_{\text{rule}, \text{visCircle}}$ met en évidence la façon dont la règle coopère avec l'entité $\mathcal{E}_{\text{visCircle}}$ qui détecte les formes circulaires, et un autre lien $\omega_{\text{rule}, \text{txt30}}$ connecte cette règle à l'entité textuelle représentant le chiffre "30". La combinaison de ces liaisons convainc ainsi que la **reconnaissance** du panneau 30 km/h découle d'une synergie tangible entre la règle symbolique "cercle + chiffre" et les observations sub-symboliques qui valident les deux prérequis.

Cette **cohabitation** symbolique–subsymbolique s'avère donc profitable pour expliquer pourquoi une règle s'active, de quelle façon elle est validée par la scène visuelle ou contextuelle, et comment cette coopération se traduit en une conclusion. Le réseau **DSL**, formé de pondérations $\{\omega_{i,j}\}$, expose explicitement la manière dont la règle " $\mathcal{E}_{\text{rule}}$ " tisse des liens avec des features visuelles ou auditives pour générer une reconnaissance de haut niveau. Il est également possible d'**extraire** des proto-règles à partir des clusters émergents du réseau. Lorsqu'un **cluster** se maintient de façon stable et regroupe certaines variables, comme des marqueurs biologiques, un symptôme et un diagnostic, on peut en déduire une **règle émergente**. Par exemple, si X, Y et Z sont élevés, la synergie avec le diagnostic D s'accroît considérablement, ce qui revient à suspecter une pathologie P. Cette forme d'**abstraction** dérive naturellement de la dynamique d'auto-organisation du **DSL** et fournit une explication concise, en reliant directement le diagnostic à un sous-réseau stable de variables médicales fortement pondérées. On franchit ainsi une étape supplémentaire vers une **explicabilité** plus intuitive, où l'on ne se contente pas de mentionner le rôle de quelques features, mais où l'on propose une **règle** ou une **raison** qui s'apparente à un énoncé symbolique, ancré dans les liaisons sub-symboliques du réseau.

1.7.5.4. Points de Vigilance : Graphes Trop Grands, Hétérogénéité Massive

Dans un **Deep Synergy Learning** de forte dimension, la **visualisation** intégrale de toutes les liaisons $\omega_{i,j}$ ou l'exploration exhaustive des chemins reliant deux entités devient inenvisageable. La taille du graphe peut atteindre plusieurs milliers ou millions de nœuds et de liens, rendant toute inspection manuelle impossible. Afin de surmonter ces limites, il est utile de mettre en place des **mécanismes** de **filtrage** ou d'**agrégation**. Le filtrage consiste à écarter les liaisons dont la pondération $\omega_{i,j}$ reste en deçà d'un **seuil** minimal, ou qui ne participent pas de manière significative à un **cluster** pivot. Cette méthode abaisse drastiquement la densité du réseau et en facilite l'interprétation. L'agrégation, quant à elle, regroupe certaines entités en **macro-nœuds**, de sorte que l'on obtienne un **graphe résumé** nettement plus petit, faisant émerger les **clusters** importants tout en préservant la cohérence des liens principaux.

Lorsque les entités du DSL recouvrent une **hétérogénéité** massive, intégrant par exemple de la vision, de l'audio, du texte et des modules symboliques, la **compréhension humaine** impose de clarifier la correspondance entre chaque entité et les données originales. Un patch d'image doit être

décrit par sa position ou sa nature dans la scène, un segment audio doit renvoyer à la phrase ou au moment précis de l'enregistrement, un concept symbolique doit être relié à son énoncé explicite. Cette **traduction** partielle est essentielle pour offrir une **explicabilité** satisfaisante. L'utilisateur comprend pourquoi une liaison $\omega_{i,j}$ est forte, car un patch visuel identifié comme "zone X" présente une **synergie élevée** avec un certain concept ou un segment audio pertinent. Il est donc indispensable d'associer à chaque entité \mathcal{E}_i un ensemble de **méta-informations** décrivant son rôle, son lien avec l'input brut et sa typologie (vision, texte, règle symbolique). Grâce à ce balisage, la **mise en récit** du réseau DSL, même de grande taille, peut s'effectuer de façon plus synthétique et plus pertinente, en se concentrant sur un sous-ensemble restreint d'arêtes fortement pondérées et sur des macro-clusters résumant la structure globale.

Conclusion

La **structure** en entités et liaisons synergiques du **Deep Synergy Learning** confère un **avantage** potentiel pour l'**explicabilité**. On peut :

- **Identifier** des clusters,
- **Exhiber** des sentiers $\mathcal{E}_i \rightarrow \mathcal{E}_j$ hautement pondérés,
- **Mettre** en avant des règles symboliques si présentes,
- **Comparer** des liens forts entre un utilisateur/contenu (dans le cas de recommandation) ou un patient/diagnostic (dans le cas médical).

Cependant, l'**explosion** du nombre d'entités et de liens peut rendre difficile la simple "lecture" de la configuration globale. On doit alors s'équiper de **techniques** de filtrage, d'**agrégation** (macro-clusters), ou de **génération** d'explications locales (chemins courts). La cohabitation possible du **symbolique** et du **sub-symbolique** aide à concrétiser des **justifications** plus claires. Ainsi, le DSL peut jouer un rôle dans la mouvance d'**Explainable AI**, à condition de manier soigneusement la **complexité** et la **variété** des entités afin de délivrer une **vision** cohérente et compréhensible à l'utilisateur final.

1.7.6. Considérations Éthiques et Réglementaires

Au-delà des enjeux techniques (scalabilité, stabilité, explicabilité), le **Deep Synergy Learning (DSL)** soulève également des questions **éthiques** et **réglementaires**. En effet, l'auto-organisation des entités au sein d'un **réseau** qui évolue sans hiérarchie prédéfinie peut introduire une forme d'**imprévisibilité** ou de **décision distribuée**. Lorsqu'il s'agit de systèmes critiques (médical, financier, militaire, etc.), il est indispensable de **responsabiliser** la conception et l'usage de tels modèles. De même, dans des applications plus courantes (recommandation, transport, logistique), la manière dont les données sont exploitées et dont les clusters se forment peut produire des **biais** ou altérer la **transparence** vis-à-vis des utilisateurs. Cette sous-section (1.7.6) examine :

- Le **risque** de perte de contrôle ou d'incompréhension du comportement d'un système auto-organisé,

- Les **biais** possibles et la nécessité de garanties contre la discrimination ou l'iniquité,
- Les **règles** et normes (RGPD, réglementations sectorielles) exigeant traçabilité et protection des données,
- Le rôle de **valeurs humaines** (sécurité, dignité, respect de la vie privée) dans la conception de DSL responsables.

1.7.6.1. Perte de Contrôle et Responsabilités

L'un des attraits majeurs du **Deep Synergy Learning (DSL)** réside dans sa **dynamique d'auto-organisation**. Les entités se regroupent en **clusters** en fonction des synergies détectées, sans nécessiter de programmation explicite pour gérer leurs interactions. Toutefois, cette souplesse peut engendrer des **comportements** difficiles à anticiper, dans la mesure où ni le concepteur, ni l'opérateur, ni le développeur n'ont directement programmé la structure finale. Dans des contextes où la **responsabilité** est cruciale, notamment en milieu médical ou bancaire, il importe de savoir **qui** assumera les conséquences si le système commet un choix discutable ou s'il engendre un dommage. Avec un réseau neuronal "classique", déjà, l'**opacité** rend l'attribution de responsabilité délicate ; dans un DSL à auto-organisation plus distribuée, le degré de complexité est plus élevé encore.

Pour pallier cette difficulté, il est envisageable de munir le système d'un **journal** d'événements et de mises à jour, permettant de **tracer** l'historique de la pondération $\omega_{i,j}$. Théoriquement, on pourrait reconstituer le "chemin causal" entre une observation et une décision, même au sein d'un graphe évolutif. Cette traçabilité demeure toutefois ardue à mettre en œuvre si l'on veut documenter en continu un réseau de grande dimension.

La loi ou les **régulations** sectorielles (ex. aéronautique, médical, bancaire) imposent souvent de **certifier** une partie du comportement du système avant son déploiement. Un DSL auto-organisé risque de "**changer**" de configuration quasi à chaque itération, rendant hasardeuse toute garantie statique. Il existe alors plusieurs stratégies. Il est possible de borner la **vitesse** d'évolution (en limitant le produit $\eta \times \tau$, ou en imposant un pas maximal) pour qu'il demeure un **voisinage** de configurations déjà validées. On peut aussi intégrer un **module** de supervision hiérarchique (voir la section 1.7.4.3) qui invalide certaines reconfigurations jugées non conformes à des normes ou des plans de sûreté. Enfin, il est envisageable d'appliquer des **tests** de robustesse, ou "stress tests", simulant divers scénarios de données pour s'assurer qu'aucune configuration délétère ne surgisse, ou du moins qu'elle reste suffisamment rare et détectable.

Ces considérations reflètent un compromis entre la **liberté** synergique, qui fait la force conceptuelle du DSL, et la **maîtrise** requise dans les secteurs critiques. Les débats sur la responsabilité sont dès lors analogues à ceux suscités par les réseaux neuronaux opacifiés, mais se complexifient en raison de la reconfiguration continue du réseau. Les utilisateurs, exploitants et autorités de tutelle doivent tenir compte de la difficulté d'isoler un instantané du DSL pour en vérifier la validité et assumer les conséquences de ses décisions évolutives.

1.7.6.2. Biais, Équité et Discrimination

Dans un **Deep Synergy Learning (DSL)** appliqué à des tâches de recommandation ou de décision, il arrive que la dynamique de formation des **clusters** reflète des **biais** préexistants dans les données, ce qui peut engendrer des effets discriminants. Imaginons un système de **scoring** pour l'attribution de crédits bancaires. Si les bases de données d'emprunteurs présentent une corrélation injustifiée entre certaines variables socio-économiques et la capacité de remboursement, le **DSL** risque de regrouper un certain sous-groupe dans un **cluster** jugé risqué, ou de l'isoler dans un espace périphérique, renforçant une forme d'exclusion. Ces liaisons faibles ou inexistantes maintiennent le sous-groupe à l'écart des opportunités, ce qui peut amplifier la ségrégation. Moins il coopère avec le reste du réseau, plus sa synergie chute, réduisant ainsi son accès équitable au crédit.

Ce phénomène d'**auto-renforcement** est potentiellement dangereux. Dans un simple réseau neuronal supervisé, des régulations ou des contraintes peuvent être imposées pour limiter l'exploitation de variables sensibles. Mais dans un **DSL**, la situation est plus complexe, car les **entités** se lient ou se séparent selon la synergie globale perçue. Un sous-groupe minoritaire, s'il n'est pas bien représenté dans les données, peut se voir marginalisé. Il convient alors d'établir des **mécanismes** préventifs ou correctifs. Il est envisageable, par exemple, de **pénaliser** l'isolation injustifiée d'un sous-groupe, en introduisant un terme additionnel dans la fonction d'énergie ou dans l'équation de mise à jour des pondérations. Ce terme peut imposer qu'un certain ensemble d'entités (associées à un attribut protégé) ne soit pas trop "distant" du reste ou qu'un cluster ne se forme pas exclusivement autour d'une caractéristique stigmatisante. Un autre moyen consiste à **rééquilibrer** les données avant l'apprentissage, afin de fournir au DSL une image plus fidèle et plus inclusive de la population, ou à injecter artificiellement des entités compensatoires, assurant que le réseau aura de quoi évaluer la synergie de manière équitable.

Dans un cadre légal, des obligations comme le **RGPD** (Règlement Général sur la Protection des Données) ou les lois anti-discrimination exigent que la décision ne dépende pas d'informations relatives à l'origine ethnique, à la religion ou à d'autres critères protégés. Le **DSL**, en créant des entités et des liaisons auto-organisées, doit donc s'assurer que certaines variables n'encoderont pas implicitement ces attributs sensibles, sous peine d'aboutir à des comportements discriminants difficilement décelables. Il est crucial de mettre en place des **audits** récurrents du réseau pour vérifier la distribution des **clusters**, observer si un groupe spécifique demeure systématiquement isolé ou sous-pondéré et, le cas échéant, forcer une mise à jour corrective. De tels audits peuvent se fonder sur l'inspection des pondérations $\{\omega_{i,j}\}$ associées à un ensemble d'entités. Ils peuvent également imposer une mesure d'équité en tant que contrainte externe, qui intervient de façon hiérarchique pour éviter la constitution de partitions inéquitables. L'objectif demeure le **contrôle** du **DSL** tout en respectant l'esprit d'**auto-organisation**, afin de combiner l'efficacité de la synergie et la nécessité de préserver une **égalité** ou une **justice** dans les décisions rendues.

1.7.6.3. Protection des Données et Droit à l'Oubli

Dans une architecture **Deep Synergy Learning (DSL)**, il arrive fréquemment que le réseau comprenne des **entités** liées à des personnes (utilisateurs, patients). Les législations comme le **RGPD** (Règlement Général sur la Protection des Données) en Europe imposent plusieurs obligations, dont la **minimisation** des données recueillies (ne conserver que l'essentiel), le **droit à**

l'oubli (un individu peut exiger la suppression de ses données personnelles) et des règles d'**anonymisation** ou de pseudonymisation pour éviter la ré-identification. Le réseau DSL, en tant que graphe auto-organisé et évolutif, doit donc intégrer des procédures garantissant que l'on puisse effacer ou rendre inopérante l'entité \mathcal{E}_i correspondant à une personne qui souhaite être oubliée.

La présence d'une entité \mathcal{E}_i dans un **Synergistic Connection Network** peut avoir des répercussions à de multiples endroits, car l'**auto-organisation** a pu former plusieurs **clusters** où \mathcal{E}_i joue un rôle. Pour satisfaire le droit à l'oubli, il faut d'abord **retracer** cette entité dans le graphe, puis la **retirer** du réseau ou la rendre totalement anonyme. Il convient de supprimer ou de neutraliser les liaisons $\omega_{i,j}$ reliées à l'entité \mathcal{E}_i . Lorsque l'on procède à cette suppression, on doit parfois **réajuster** la **synergie** locale afin que les entités \mathcal{E}_j anciennement liées à \mathcal{E}_i ne se retrouvent pas avec des pondérations invalides. L'opération de "déréférencement" peut donc influencer sur l'équation d'évolution, qui doit détecter et ignorer toute trace associée à \mathcal{E}_i . L'objectif est d'éviter que des **résidus** dans la structure du réseau ne permettent de re-identifier l'utilisateur a posteriori, contrevenant ainsi aux principes de confidentialité exigés par la loi.

Le RGPD recommande également la **pseudonymisation** et la **minimisation** des informations stockées, ce qui affecte la façon dont les entités DSL sont construites. Il est parfois nécessaire de ne stocker, dans une entité \mathcal{E}_i , qu'une version agrégée (par exemple, localisation au niveau d'une ville plutôt que d'une rue) et d'exclure tout identifiant direct pour se conformer aux dispositions légales. Si le réseau DSL garde un trop haut niveau de précision, il peut être considéré comme non conforme à la réglementation. Une approche consiste à instaurer un **processus** d'anonymisation périodique. Après un certain temps, on fusionne les entités trop granulaires ou on supprime certains détails, réduisant la granularité tout en préservant l'essence des **synergies** globales. Cette pratique garantit une protection de la vie privée tout en permettant au DSL de continuer à fonctionner de manière auto-organisée et réactive, mais sur la base de données moins sensibles ou déjà anonymisées.

1.7.6.4. Transparence, Explicabilité et Contrôle

Dans certains contextes réglementés, comme le **RGPD** ou des lois spécifiques à un domaine (médical, bancaire, etc.), les usagers ou patients ont la faculté d'exiger une **explication** sur la décision qui les concerne. Le **Deep Synergy Learning (DSL)**, grâce à ses **clusters** et à ses **liaisons** plus faciles à interpréter qu'un pipeline neuronal opaque, peut répondre à cette exigence. Il demeure toutefois nécessaire de s'assurer que l'évolution du réseau (les pondérations $\omega_{i,j}$ dans le Synergistic Connection Network) demeure **traçable** dans le temps. Pour cela, on peut tenir un **historique** des principales mises à jour ou conserver des **snapshots** à intervalles réguliers, afin de justifier a posteriori comment tel ou tel sous-réseau s'est formé.

Une **infrastructure** d'explication (voir section 1.7.5) doit de surcroît synthétiser le résultat de la **dynamique** auto-organisée. Il ne s'agit pas nécessairement d'exposer l'entièreté des liens $\omega_{i,j}$, mais d'extraire des **chemins** ou des **clusters** pivot permettant de reconstituer le **raisonnement local**. Une difficulté apparaît lorsque le système évolue de manière rapide ou chaotique. L'explication peut alors varier significativement d'une itération à l'autre, compromettant la cohérence globale. Des **mécanismes** de stabilisation ou de **pas** d'évolution limités (contrôle de η

et de τ , par exemple) peuvent aider à préserver une relative continuité, rendant les explications plus consistantes dans le temps.

Pour de nombreuses **applications** sensibles, la législation ou la pratique courante impose un **contrôle humain** dans la boucle. Conformément à un principe d’audit ou de validation, un expert (médecin, analyste, manager) examine la configuration actuelle du DSL et doit pouvoir **accepter**, **refuser** ou **rectifier** certaines décisions. L’**interface** qui présente les clusters, leurs pondérations dominantes et l’historique succinct de leur formation doit donc être la plus claire possible, en hiérarchisant l’information. L’expert peut alors intervenir en imposant une pénalisation sur des liaisons considérées comme non conformes (sécurité, éthique, cohérence métier). Cette démarche articule la liberté **auto-organisée** du DSL (qui adapte continûment ses liaisons) avec un **pilotage** humain, assurant un respect des **cadres légaux** et éthiques, tout en tirant parti de la plasticité et de l’auto-organisation qu’offre le Deep Synergy Learning.

Conclusion

Les **considérations éthiques et réglementaires** forment un **pilier** incontournable lors du déploiement de **systèmes auto-organisés** comme le DSL. On peut résumer les **grandes lignes** de vigilance :

- **Responsabilité et traçabilité** : éviter qu’un système auto-adaptatif devienne totalement opaque quant à ses décisions ou conséquences, mettre en place un journal d’événements, un module de supervision.
- **Biais et équité** : surveiller la formation de clusters discriminants, adapter la dynamique pour prévenir les exclusions ou dominations injustes.
- **Protection des données** : garantir l’anonymisation, la pseudonymisation, et le droit à l’oubli dans un graphe évolutif (ce qui n’est pas trivial techniquement).
- **Transparence et droit à l’explication** : mettre en place des mécanismes d’explicabilité (voir 1.7.5) assurant que l’utilisateur comprend la logique du DSL.

Au final, l’**originalité** du DSL ne saurait dispenser d’une **responsabilisation** solide. La recherche sur l’IA dite “responsable” ou “de confiance” s’applique tout autant (voire davantage) à un modèle **auto-organisé** qui, par définition, échappe aux schémas de contrôle univoque. Les prochaines sections (et développements futurs) devront approfondir les outils et protocoles pour faire du DSL une **technologie** à la fois puissante, adaptable, et conforme aux **valeurs** et **règles** en vigueur.

1.7.7. Comparaisons Expérimentales avec d’Autres Approches

Dans les sections précédentes (1.7.1 à 1.7.6), nous avons exploré les **défis** et **contraintes** du **Deep Synergy Learning (DSL)** en matière de **complexité**, de **qualité des données**, de **stabilité**, d’**explicabilité**, et d’**éthique**. Pour renforcer la **pertinence** de ce paradigme en pratique, il est essentiel de réaliser des **comparaisons expérimentales** rigoureuses avec des méthodes existantes. Ces évaluations incluent les **réseaux neuronaux profonds** (CNN, RNN, Transformers), les

méthodes de clustering (k-means, DBSCAN), les **algorithmes d'optimisation** (ordonnancement) et les **approches neuro-symboliques** traditionnelles. Cette sous-section (1.7.7) aborde :

- Les **critères** de comparaison essentiels (performances quantitatives, robustesse, adaptativité...),
- Les **protocoles** expérimentaux pour confronter le DSL à d'autres techniques (tests sur données stationnaires et non stationnaires, analyses de complexité, etc.),
- Les **scénarios** où le DSL a un net avantage (auto-organisation, peu de supervision, scénarios évolutifs),
- Les **limites** et enseignements que de tels benchmarks peuvent révéler, guidant l'amélioration du DSL.

1.7.7.1. Choisir les Critères de Comparaison

Dans la plupart des tâches d'apprentissage ou de décision (vision, audio, recommandation, diagnostic médical), on évalue la **précision** ou des métriques dérivées comme l'**accuracy**, la **F1-score** ou l'**AUC**. Lorsque l'on souhaite comparer un **Deep Synergy Learning (DSL)** avec, par exemple, un CNN ou un Transformer, on peut se borner à mesurer son **taux de reconnaissance** (dans un scénario de classification d'images) ou la **sensibilité** et la **spécificité** (dans le cadre d'un diagnostic médical). Toutefois, ces **scores** globaux ne reflètent pas nécessairement la **richesse** de la structure auto-organisée du DSL, qui excelle par sa **plasticité** et son adaptation continue, sans exiger nécessairement de labels massifs. Il est donc essentiel, lors de l'évaluation, d'incorporer des critères tenant compte de la **dynamique** de l'auto-organisation.

Le **DSL** se veut **adaptatif** en présence de bruit, de pannes de capteurs ou de changements de distribution (concept drift). Pour mettre en évidence cette capacité, il est concevable de concevoir des expériences où la distribution des données varie progressivement ou subitement en cours d'apprentissage. Dans un cas classique, un **réseau neuronal** (CNN, RNN) va souvent exiger une **séance de re-fine-tuning** global, tandis que le DSL doit pouvoir **réajuster** localement ses liaisons $\{\omega_{i,j}\}$. On peut alors comparer la **vitesse de réadaptation** et la **qualité** de la reprise de performance. On mesure, par exemple, la **chute** temporaire de précision lorsqu'un nouveau type de données apparaît, ainsi que le **temps** qu'il faut au modèle pour retrouver un niveau satisfaisant. Le DSL a pour atout de pouvoir réorganiser ses clusters et de créer ou rompre des liens en fonction des synergies locales, ce qui peut se révéler plus rapide et moins coûteux qu'un re-entraînement complet.

La **taille** du modèle (nombre de paramètres ou de liaisons actives) et la **charge** de calcul (temps CPU, mémoire GPU) comptent également parmi les indicateurs déterminants. Un DSL peut rester **parcimonieux** s'il opère une **sparsification** rigoureuse de la matrice ω . Si, au contraire, on ne limite pas la prolifération des liens, la structure risque de gonfler et de devenir onéreuse en ressources. L'évaluation doit inclure la **scalabilité** et l'**efficacité**. Il est crucial de vérifier si, face à un grand nombre d'entités, le **DSL** maintient une densité de liens contrôlée (voir section 1.7.1) et s'adapte aux scénarios évolutifs tels que les pannes ou les modifications de distribution. On

évalue également les **coûts de mise à jour** et de **mémoire**, afin de garantir une gestion efficace des ressources.

1.7.7.2. Protocoles Expérimentaux Envisageables

Pour **évaluer** la pertinence d'un **Deep Synergy Learning (DSL)** dans des scénarios variés, on peut le confronter à des bases de données classiques ou à des cas pratiques bien établis :

- Des ensembles **d'images** (MNIST, CIFAR, éventuellement un sous-ensemble d'ImageNet) pour évaluer la classification ou la détection, en analysant la **précision** obtenue par l'auto-organisation visuelle des entités (patches, features, etc.) face à un CNN ou un Transformer.
- Des ensembles **audio** (LibriSpeech, UrbanSound8K) pour tester la reconnaissance de sons ou la transcription de parole, où le DSL formerait des entités "segment audio" et "caractéristiques acoustiques", comparées à un système RNN ou CNN entraîné de façon traditionnelle.
- Des **bases de recommandation** (MovieLens, Amazon reviews) pour mesurer la qualité des suggestions formulées par le DSL ; on confronte alors la structure auto-organisée (clusters d'utilisateurs, de contenus, etc.) aux méthodes de factorisation matricielle ou aux modèles de collaboration neuronaux, en examinant le **Recall@K**, la **Precision@K** ou la **NDCG**.
- Des **données médicales** (ensembles de pathologies, EHR) afin de juger la capacité du DSL à découvrir des clusters pathologiques ou à émettre des hypothèses de diagnostic, comparativement à un modèle supervisé (réseau profond, arbre d'ensemble) ou semi-supervisé.
- Des **scénarios robotiques** utilisant des simulateurs comme **Gazebo** ou **PyBullet** permettent d'évaluer la coordination sensorimotrice. Un **DSL** connecte capteurs et effecteurs, testant sa **robustesse** face aux variations des conditions et à la détection de pannes.

Dans chaque domaine, les **résultats** finaux tels que le taux de reconnaissance, le score de recommandation ou le F1-score diagnostique sont comparés aux performances de modèles **classiques** comme les **CNN**, **RNN**, **Transformers**, **K-means** ou la **factorisation matricielle**. L'évaluation inclut également la **taille du modèle** en termes de paramètres et de liaisons, la **complexité computationnelle**, ainsi que la **vitesse d'inférence**.

Le **DSL** se veut également **adaptatif** en cas de changements de distribution (concept drift), de pannes de capteur, de conditions altérées. Pour l'exercer à cette fin, il est utile de mettre en œuvre des **scénarios** où les données subissent un glissement graduel ou une modification brutale (apparition de nouvelles classes, hausse du bruit, changement d'éclairage, etc.). Sur la timeline $\{t_0, \dots, t_f\}$, on évalue deux aspects. On observe comment un réseau neuronal classique comme un CNN ou un RNN réagit face au drift des données, souvent en nécessitant un re-fine-tuning ou en subissant une perte partielle d'acquis menant à un catastrophic forgetting. On analyse également comment le **DSL**, sans recourir à un réapprentissage massif, ajuste localement ses liaisons $\omega_{i,j}(t)$ afin de s'adapter aux nouvelles conditions.

On observe la **performance** en continu (taux de classification, AUC, indice de recommandation) et on étudie la **vitesse** de réadaptation. Le **DSL** devrait, en principe, reporter des ajustements plus localisés et moins coûteux qu'un re-entraînement total, tout en préservant une cohérence globale. Ceci permet de vérifier empiriquement que la **robustesse** et la **plasticité** conférées par la mise à jour distribuée des pondérations ω confèrent un avantage aux méthodes d'auto-organisation, surtout dans des domaines où les données ou le contexte varient régulièrement.

1.7.7.3. Scénarios où le DSL a un Avantage Notable

Le **Deep Synergy Learning (DSL)** se distingue particulièrement dans des environnements où l'on dispose de **peu ou pas de labels** (section 1.5.5) ou bien où la supervision est coûteuse et incomplète. En mode **non supervisé** ou faiblement supervisé, un réseau neuronal classique (tel qu'un CNN) peine à extraire des structures cohérentes, car il manque d'exemples étiquetés pour guider l'apprentissage. Le **DSL**, au contraire, s'appuie sur sa **fonction de synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$ pour **auto-organiser** les entités, sans nécessiter un volumineux jeu de labels. Les **clusters** émergent alors de la simple coopération (ou co-occurrence) observée entre entités, ce qui rend le réseau particulièrement apte aux tâches d'**exploration** et de **regroupement** de données hétérogènes.

Un autre atout du DSL s'observe en **multi-modalité hétérogène** (section 1.5.2), où l'on combine image, audio, texte, voire des flux sensoriels divers. Les approches traditionnelles imposent souvent un **pipeline** de fusion rigide (concatenation tardive, attention inter-modale) dépendant de choix d'architecture prédéfinis. Le **DSL** adopte une approche plus flexible où chaque entité \mathcal{E}_i (visuelle, auditive, etc.) établit des **liaisons** $\omega_{i,j}$ en fonction de la **synergie** détectée. Ces connexions émergent et évoluent sans qu'un schéma fixe ne soit imposé, permettant une organisation dynamique du réseau. Cette structure évite la planification d'un module de "fusion" spécifique ; la mise en réseau est régie par la **dynamique** locale des pondérations, autorisant des clusters multimodaux à émerger naturellement.

Lorsque les **flux** de données sont **dynamiques** et sujets à des **pannes** ou à des transformations (sections 1.5.4 et 1.7.2.3), le DSL illustre également sa supériorité. Un capteur défectueux voit ses liaisons ω chuter, car la synergie avec le reste du système devient faible. Le réseau se réorganise alors sans qu'un nouvel entraînement global ne soit requis, contrairement à un réseau neuronal classique où une modification durable de la distribution (panne de capteur, concept drift) nécessite un **fine-tuning** ou un recalibrage important.

L'**explicabilité** du DSL revêt aussi un intérêt clé (sections 1.7.5 et 1.5.6). Les **clusters** ou les chemins de **synergie** identifiés entre entités procurent des explications plus détaillées qu'une simple "heatmap" d'attention. Dans un réseau traditionnel, la structure interne (couches, neurones) reste souvent opaque, malgré certaines techniques de visualisation de filtres. Le DSL, en attribuant des **identités** distinctes à ses nœuds et en laissant les liens $\omega_{i,j}$ témoigner de leur degré de coopération, permet une **compréhension** plus directe de la raison pour laquelle tel sous-réseau est né, ou tel couple $(\mathcal{E}_i, \mathcal{E}_j)$ a vu sa liaison renforcée.

Dans l'ensemble, une **comparaison expérimentale** (section 1.7.7.2) révélera que, face à un CNN ou à un Transformer, le DSL excelle dans des conditions de **data faible** ou de **distribution changeante**, ou encore quand l'on souhaite un **rendement** sémantique mieux explicité par un

graphe de clusters. Ces propriétés séduisent dans des scénarios où l'on doit adapter le système en continu, respecter des contraintes d'équité, ou justifier les décisions auprès d'utilisateurs ou d'experts.

1.7.7.4. Limites et Enseignements Possibles

Lors de l'évaluation d'un **Deep Synergy Learning (DSL)** à large échelle, on peut constater que la structure du réseau, qui autorise jusqu'à $O(n^2)$ liaisons dans le cas binaire, devient difficile à manipuler si les entités sont très nombreuses. Il peut être nécessaire d'imposer une **sparsification** stricte (par exemple, ne maintenir qu'un voisinage restreint pour chaque entité) ou de recourir à un **algorithme local** qui ne calcule pas la synergie entre toutes les paires d'entités, mais seulement entre celles jugées proches dans un espace de représentation. Les expérimentations comparatives, que ce soit en temps de calcul ou en mémoire GPU, permettent d'identifier le **seuil** au-delà duquel le DSL se montre moins performant que d'autres architectures classiques (CNN, RNN, Transformer) optimisées pour le passage à l'échelle sur des supports matériels massivement parallèles.

Par ailleurs, si l'on dispose d'un **jeu de données étiquetées** gigantesque (par exemple la totalité d'ImageNet, avec plus de 14 millions d'images), un réseau supervisé standard tel qu'un CNN ou un Transformer minutieusement entraîné peut atteindre d'excellentes performances de classification top-n. Le **DSL**, plus orienté vers l'**auto-organisation** non supervisée ou faiblement supervisée, pourrait dans ce cas demeurer en retrait si l'on n'exploite pas pleinement toute la supervision à disposition. On peut envisager des versions **supervisées** ou **semi-supervisées** du **DSL** où la **synergie** est partiellement orientée par des labels. Toutefois, une question demeure ouverte. Ces labels seront-ils suffisants pour rivaliser avec un **CNN** entraîné par descente de gradient pure, en particulier lorsque le volume d'annotations est massif ? Ces expériences permettront de clarifier dans quelle mesure le DSL parvient à tirer parti d'une supervision large sans renoncer à sa plasticité et à sa modularité intrinsèque.

Ces limites et enseignements soulignent la nécessité de concevoir des **protocoles hybrides**, dans lesquels on combine la liberté d'auto-organisation du DSL (favorable à l'adaptation continue ou à la gestion de scénarios peu labellisés) avec la force d'un apprentissage supervisé classique lorsqu'un dataset étiqueté d'ampleur est disponible. C'est dans cet espace de compromis que réside probablement l'avenir des approches synergiques à l'échelle industrielle.

Conclusion

La **validation expérimentale** du **DSL** face aux approches dominantes, comme les réseaux neuronaux profonds, le clustering ou les méthodes d'optimisation, est essentielle pour **quantifier** sa performance, sa robustesse et sa capacité d'adaptation. Elle permet également d'**identifier** ses atouts, notamment l'apprentissage sans label massif, la gestion du drift et l'intégration de modalités hétérogènes. Enfin, elle aide à **révéler** ses limites, telles que le coût en calcul, la nécessité de mesures de synergie fiables et les difficultés sur des tâches fortement supervisées.

Des **benchmarks** variés (vision, audio, recommandation, diagnostic, robotique) permettront de **cartographier** où le DSL s'impose et où il doit s'allier à d'autres techniques (CNN pré-entraîné pour extraire des features, par exemple). Finalement, ce travail de confrontation expérimentale, mené en parallèle des réflexions éthiques (1.7.6) et de la recherche sur l'optimisation (1.7.3), consolide la position du DSL comme un **paradigme** original — potentiellement complémentaire, voire compétitif — dans le paysage actuel de l'apprentissage automatique.

1.8. Positionnement du DSL dans l'Évolution de l'IA

Après avoir étudié les fondements du **Deep Synergy Learning (DSL)** et ses multiples applications (sections 1.1 à 1.6), ainsi que les défis et contraintes (section 1.7), il est opportun de s'interroger sur la **place** qu'occupe le DSL dans l'**évolution** de l'intelligence artificielle. L'IA contemporaine repose sur deux grands paradigmes. D'un côté, l'**IA sub-symbolique**, incarnée par les réseaux neuronaux et les méthodes statistiques, s'est imposée avec l'essor de l'apprentissage profond. De l'autre, l'**IA symbolique**, issue des systèmes experts et de la logique formelle, offre une meilleure interprétabilité et une manipulation plus explicite sous forme de langage structuré.

Le **DSL**, en proposant une **auto-organisation** et une **dynamique** des liens fondée sur la synergie, apporte un **paradigme** susceptible de **réconcilier** certains points de friction entre ces approches. On peut se demander :

- S'il offre un **point** de convergence entre l'**IA symbolique** et l'**IA sub-symbolique**,
- Comment il peut **coexister** ou se **substituer** à l'apprentissage profond,
- Dans quelle mesure le **renforcement** (RL) ou la **logique** se fondent dans une trame DSL,
- Le **rôle** d'éléments comme la **mémoire** et l'**attention** dans une architecture auto-organisée,
- Les **tendances** futures, notamment vers une IA plus "forte" ou consciente,
- L'**impact** interdisciplinaire, en particulier avec les neurosciences,
- Et, de manière générale, comment le DSL s'inscrit dans la **progression** de l'IA moderne et ce qu'il peut apporter de nouveau.

Cette section (1.8) aborde ainsi le **positionnement** du DSL dans le paysage AI actuel et futur, regroupant sept sous-points :

- **IA Symbolique vs IA Sub-symbolique : Intégration Potentielle** (1.8.1)
- **DSL et Apprentissage Profond : Collaboration ou Substitution ?** (1.8.2)
- **Approches Hybrides : DSL, RL (Reinforcement Learning) et Logique** (1.8.3)
- **Rôle de la Mémoire et de l'Attention dans le DSL** (1.8.4)
- **Tendances Futures : Vers une IA Forte, Consciente ?** (1.8.5)
- **Effet Sur la Recherche Interdisciplinaire** (1.8.6)
- **Exemple de Convergence : DSL & Neurosciences** (1.8.7)

Nous commencerons par revisiter (1.8.1) la vieille dichotomie IA symbolique vs sub-symbolique, pour voir en quoi le **DSL** pourrait établir des passerelles novatrices.

1.8.1. IA Symbolique vs IA Sub-symbolique : Intégration Potentielle

L'**histoire** de l'IA est souvent présentée comme marquée par un **dualisme** :

- **L'IA symbolique**, qui manipule des règles, des faits, des ontologies dans un langage logique ou pseudo-logique (ex. Prolog, systèmes experts). Elle est réputée plus “interprétable” et plus “rigoureuse” dans le raisonnement, mais manque de flexibilité face aux données massives et bruitées.
- **L'IA sub-symbolique**, typiquement les réseaux neuronaux et méthodes statistiques, qui excellent en **apprentissage** (vision, NLP, etc.) en profitant de nombreux exemples, mais peinent à manipuler des structures conceptuelles complexes ou des règles explicites.

1.8.1.1. Vers une IA Hybride ou Neuro-Symbolique

Au cours de la dernière décennie, un nombre croissant de recherches s'est attelé à **combiner** les approches symboliques (issues de la logique, des règles ou des ontologies) et les méthodes sub-symboliques (neuronales ou connexionnistes). Les travaux dits **neuro-symboliques** cherchent, par exemple, à implanter des **règles logiques** au sein d'un réseau neuronal ou, à l'inverse, à extraire des **règles** symboliques depuis les *features* d'un modèle déjà entraîné. L'objectif est de combiner la **puissance** des méthodes sub-symboliques, capables de traiter de grands volumes de données complexes, avec la **rigueur** et la **lisibilité** des approches symboliques, qui facilitent le raisonnement et l'explication.

Ces approches neuro-symboliques s'organisent souvent sous la forme de **systèmes hybrides**, où l'on place côte à côte une **composante** consacrée au traitement sub-symbolique (réseau neuronal) et une **composante** symbolique (règles ou moteur logique). Un **pont** dédié est alors chargé de faire correspondre ces deux volets, ce qui peut induire une architecture “collée” au sens où la dimension symbolique reste nettement délimitée de la dimension connexionniste. La **véritable** intégration, où le réseau lui-même serait libre d'accueillir à la fois des **concepts logiques** et des **signaux** bruts en adaptant sa topologie, demeure un défi qui n'est pas entièrement résolu dans la plupart de ces schémas.

Le **Deep Synergy Learning (DSL)** entend précisément répondre à cette problématique, car il propose une **dynamique** où la structure se reconfigure **auto-organisationnellement**, sans imposer de cloison fixe entre l'aspect symbolique et l'aspect sub-symbolique. Les **entités** \mathcal{E}_i peuvent incarner aussi bien un **flux** sensoriel (image, audio), qu'un **concept** symbolique, ou une **règle**. Les liaisons $\omega_{i,j}$ se créent ou se coupent selon la **synergie** détectée. Un nœud représentant une règle logique pourra donc se lier plus fortement à un patch d'image ou un module audio dès lors qu'il y trouve un **gain** mutuel. De cette manière, la frontière traditionnelle entre le bloc **symbolique** et le bloc **neuronal** cède la place à un **réseau** unifié, où chaque composante est traitée comme une entité à part entière, susceptible de coopérer librement pour faire émerger des **clusters** ou des **schémas** de raisonnement. Ce fonctionnement souligne l'ambition du DSL de faciliter la rencontre fluide entre la **rigueur** des modèles logiques et la **flexibilité** d'apprentissage des méthodes sub-symboliques.

1.8.1.2. Rôle du DSL : Auto-Organisation Unifiant Symbolique et Sub-Symbolique

Le **Deep Synergy Learning (DSL)** propose une **intégration** fluide entre les entités **sub-symboliques** (capteurs, features, embeddings neuronaux) et les **composantes symboliques** (règles, concepts logiques, ontologies). Dans la dynamique du DSL, les **liaisons** $\omega_{i,j}$ se créent ou se rompent en fonction de la **synergie** perçue (voir la section 1.4.5 pour la mise à jour itérative). Au lieu de séparer la logique dans un module détaché, on traite chaque **règle** ou **concept** comme un **nœud** $\mathcal{E}_k^{(\text{symb})}$ dans le **Synergistic Connection Network**. La coopération entre une règle symbolique $\mathcal{E}_k^{(\text{symb})}$ et une entité sub-symbolique \mathcal{E}_i (par exemple un capteur) repose sur une pondération $\omega_{\text{symb},i}$ qui s'adapte selon la synergie. Si la règle se vérifie fréquemment grâce aux données de ce capteur, la liaison se renforce, tandis que dans le cas contraire, elle s'affaiblit progressivement.

Dans un **environnement** robotique, plusieurs **règles** logiques \mathcal{R}_k peuvent être définies pour assurer la sécurité et la performance des machines. Par exemple, une règle pourrait stipuler que si la **tension moteur** dépasse **5V** et que la **température moteur** excède **80°C**, alors une alerte de surchauffe doit être déclenchée.

Dans le cadre du **DSL**, une telle règle est modélisée sous forme d'une **entité symbolique** $\mathcal{E}_k^{(\text{symb})}$. Parallèlement, les différentes **entités sub-symboliques** représentent les capteurs physiques, comme la mesure de la tension et de la température. La **synergie** entre ces éléments se traduit par les pondérations $\omega_{\text{tension}, \text{règle}}$ et $\omega_{\text{temp}, \text{règle}}$, qui augmentent si les relevés des capteurs confirment régulièrement la condition de surchauffe.

Sans imposer une structure figée, le **réseau** établit un **micro-réseau** reliant de façon dynamique la règle symbolique aux capteurs pertinents. Si la synergie persiste dans le temps, un **cluster** stable se forme, associant la règle “surchauffe moteur” aux données sensorimotrices qui la valident. Cette organisation **auto-adaptative** permet de lier les règles aux données pertinentes sans supervision explicite.

Cette approche évite le recours à une **passerelle** rigide entre la logique formelle et les signaux bruts. La **topologie globale** du **DSL** se **réajuste** en fonction de la pertinence observée. Si la règle \mathcal{R}_k devient moins utile ou si de nouveaux **capteurs** plus adaptés émergent, les **pondérations** s'adaptent automatiquement. On obtient ainsi une **auto-organisation** où la dimension **symbolique** est pleinement intégrée à la dynamique décrite en 1.4.5, donnant un **substrat** unifié pour l'IA hybride.

1.8.1.3. Avantages et Limites

Le fait d'intégrer la **dimension symbolique** directement dans un **Deep Synergy Learning (DSL)** offre divers **avantages**. D'abord, la **logique** s'intègre directement à l'**architecture auto-organisée**. Il n'est plus nécessaire de superposer un **module de règles** indépendant à un **module sub-symbolique**, comme un réseau neuronal. Au sein du DSL, chaque règle \mathcal{R}_k devient une entité $\mathcal{E}_k^{(\text{symb})}$, et la **pondération** $\omega_{\text{symb},i}$ traduit la coopération effective entre cette règle et n'importe quelle entité sub-symbolique \mathcal{E}_i . Cette façon de procéder rend possible l'**apprentissage** continu de l'utilité de la règle, puisque la pondération se régule en fonction de la “valeur ajoutée” détectée. Un autre **bénéfice** se manifeste en matière d'**explicabilité**. Il devient plus facile d'identifier un

cluster reliant une règle \mathcal{R}_k à des **entités perceptives** $\mathcal{E}_p, \mathcal{E}_q, \dots$, permettant ainsi de justifier une décision. On obtient dès lors une **transparence** plus claire qu’avec un système classiquement scindé où la logique reste à part.

Cependant, cette intégration comporte également des **limites** notables. Une première difficulté tient à la **quantification** de la synergie $S(\mathcal{E}_i, \mathcal{E}_k^{(\text{symb})})$ quand l’une des entités représente une **règle** ou un **concept** logique. Il s’agit de définir mathématiquement la notion de “gain” associé à l’activation conjointe d’une règle symbolique et d’un flux sub-symbolique. Ce n’est pas toujours trivial, surtout si la règle mélange des conditions logiques complexes. Par ailleurs, plus on introduit de **règles**, plus la combinatoire peut **exploser**, le réseau devant gérer de multiples $\omega_{(\text{symb}),i}$. La mise à jour itérative de ces liaisons doit alors tenir compte des possibles **contradictions** entre règles (ex. deux règles incompatibles ne sauraient être activées simultanément), ce qui alourdit la dynamique ω et nécessite des **mécanismes** de gestion de conflits ou de contradictions logiques.

Ces considérations exigent une **régulation** attentive, par exemple via des **termes** de parcimonie (limitation du nombre total de liaisons actives), des **pénalités** imposées aux liaisons incompatibles, ou un **contrôle** de la vitesse de croissance de certaines pondérations. Un tel encadrement évite l’explosion combinatoire de la structure et aide à résoudre de manière cohérente les potentielles contradictions entre règles symboliques. Malgré ces contraintes, la promesse demeure. Le **DSL** offre un cadre plus **organique**, où les **concepts** logiques et les **descripteurs** sub-symboliques coexistent de manière adaptative.

1.8.1.4. Perspectives sur l’Intégration Totale

Dans les modèles traditionnels, une **dichotomie** claire oppose l’IA **symbolique**, qui repose sur des règles logiques et la manipulation de symboles, et l’IA **sub-symbolique**, qui exploite les réseaux neuronaux et les méthodes connexionnistes pour traiter les données brutes par apprentissage statistique. Le **Deep Synergy Learning (DSL)** aspire à dépasser cette fragmentation en poussant l’**auto-organisation** à un niveau plus global. Les **entités symboliques** – vues comme de véritables “super-neurones” capables d’inférences logiques internes – coexistent avec des **entités sub-symboliques** qui extraient, à partir de données brutes, des features ou des embeddings. Toutes ces entités demeurent dans un **même** Synergistic Connection Network, au sein duquel les liens $\omega_{i,j}$ se créent, se renforcent ou s’affaiblissent en fonction de la synergie effectivement observée.

Lorsqu’une **règle symbolique** apporte un “gain” (par exemple, elle s’active souvent conjointement à un certain pattern perceptif), la liaison entre cette règle et les entités perceptives impliquées croît. Inversement, si une règle n’est presque jamais appliquée ou se trouve contredite par les observations, ses liens se dissolvent. De la même manière, certains **macro-clusters** voient le jour, rassemblant des **concepts** sémantiques et des **patches** perceptuels (image, audio, etc.) autour d’une **synergie** élevée. Un tel réseau ne fixe pas de **cloisonnement** rigide entre la dimension **symbolique** et la dimension **sub-symbolique**. Les deux coexistent et interagissent librement, se distinguant uniquement par leurs **rôles** respectifs au sein de l’architecture auto-organisée.

Cette intégration n’est pas strictement spéculative et trace la voie d’une **IA cognitive** plus étendue, qui mêlerait la **perception** (apprentissage sub-symbolique) et le **raisonnement** (entités logiques) au sein d’un seul et même **réseau** en évolution. Les applications potentielles sont nombreuses. Un système de **vision** peut enrichir son raisonnement en exploitant des règles symboliques, qu’elles soient définies ou émergentes. De même, un moteur **logique** bénéficie de la synergie avec les flux

sensoriels, renforçant ainsi l'interprétation et l'adaptabilité du réseau. Le **DSL** rend possible la création d'un **substrat** où, au fil de l'apprentissage, les connexions s'**enrichissent**, des **règles** naissent ou se dissolvent, et des **clusters** conceptuels émergent, sans qu'on doive prescrire a priori leur structuration. Le réseau devient un "lieu" de coopération continue entre la dimension **symbolique** et la dimension **connexionniste**, accréditant l'idée d'une IA vraiment **hybride**, capable de percevoir et de raisonner sans barrière artificielle.

1.8.2. DSL et Apprentissage Profond : Collaboration ou Substitution ?

Alors que le **Deep Learning** (au sens des réseaux de neurones profonds) occupe aujourd'hui une place dominante dans la recherche et les applications de l'IA, on peut se demander si le **Deep Synergy Learning (DSL)** vient s'y substituer ou, au contraire, s'il peut s'**allier** à ces techniques. En effet, le **DSL** propose une vision **auto-organisée** où les entités d'information découvrent et renforcent leurs liens synergiques, tandis que l'apprentissage profond repose typiquement sur des **architectures hiérarchiques** (CNN, RNN, Transformers) entraînées par **descente de gradient** sur un ensemble de données massif. Cette sous-section (1.8.2) explore la relation entre ces deux paradigmes en analysant s'ils sont en **concurrence**, **complémentaires**, ou destinés à une **convergence** progressive.

1.8.2.1. Les Forces Reconnues du Deep Learning

Depuis plus d'une décennie, l'**apprentissage profond** a accompli des progrès spectaculaires dans plusieurs domaines clés de l'intelligence artificielle. En **vision artificielle**, on constate sa capacité à exceller dans la **classification** d'images ou la **détection** d'objets, tandis qu'en **traitement du langage naturel**, il domine pour les tâches de **traduction**, de **question-réponse** ou de **résumé** automatique. On retrouve également des succès notables en **analyse audio**, en particulier dans la **reconnaissance vocale**, et dans des environnements de **jeux** comme ceux explorés par AlphaGo ou AlphaZero. Ces avancées reposent en premier lieu sur la **puissance** de l'apprentissage supervisé lorsqu'il est alimenté par de larges volumes de données annotées. La deuxième composante décisive réside dans les **architectures** spécialisées du deep learning (réseaux de neurones convolutifs en vision, Transformers en langage, etc.), qui exploitent la structure des données pour atteindre des performances remarquables. Enfin, ces progrès sont intimement liés au développement accéléré des **librairies logicielles** (TensorFlow, PyTorch...) et du **matériel** (GPU, TPU) dédiés aux calculs massivement parallélisés.

Malgré ces forces, le deep learning fait souvent l'objet de **critiques** qui mettent en évidence sa dépendance à une supervision très abondante, et sa difficulté à se réorienter lorsque la tâche ou le domaine change. Son cadre global demeure souvent **rigide**. Lorsqu'une nouvelle tâche survient, il faut procéder à une **reconfiguration** ou un **fine-tuning** spécifique, ce qui complique l'adaptation dans un environnement en perpétuelle évolution. Les réseaux neuronaux profonds manquent également de **plasticité** et de **résilience** lorsque des pannes de capteur ou des modifications du contexte perturbent la distribution des données ; la moindre altération exige un nouvel entraînement ou une nouvelle adaptation globale. S'y ajoute la question de la **transparence**. Le caractère **boîte noire** des modèles pose un défi majeur pour l'explicabilité. Il devient difficile de comprendre

précisément l'origine d'une prédiction ou d'une décision, surtout lorsque le modèle comporte des dizaines de couches et des millions, voire des milliards, de paramètres.

1.8.2.2. Le DSL comme Complément : Auto-Organisation et Adaptativité

Le **Deep Synergy Learning (DSL)** repose sur une **auto-organisation** distribuée. Des **entités**, qu'il s'agisse de *features* neuronales, de blocs d'apprentissage ou de règles symboliques, établissent des connexions dynamiques à travers des **pondérations synergiques** $\omega_{i,j}$. Ces pondérations évoluent en continu selon une loi d'**adaptation** locale, autorisant l'insertion ou la suppression d'entités sans nécessiter une refonte ou un réentraînement global du réseau. Ce mécanisme confère une **plasticité** et une **résilience** particulières, notamment en cas de nouveaux flux de données ou de variations contextuelles.

Il est aisé d'envisager un **scénario** de **collaboration** entre le deep learning traditionnel et le DSL. Par exemple, on peut employer un **CNN** (ou un **Transformer**) pour extraire des **représentations profondes** à partir d'images, de signaux audio ou de séquences textuelles. Ces représentations (vecteurs de *features*) se traduisent alors en un **ensemble** d'entités $\{\mathcal{E}_{\text{CNN}}\}$ reliées entre elles ou avec d'autres entités par des pondérations ω . Le **DSL** assure une **auto-organisation** dynamique en identifiant les **synergies** entre divers blocs de *features* ou entre *features* et règles symboliques. Il ajuste les **clusters** émergents en fonction de leur pertinence, qu'ils soient multimodaux, sémantiques ou contextuels. L'ajustement local des $\omega_{i,j}$ s'effectue de manière à ce que, si deux groupes de *features* se complètent fortement, un **macro-cluster** puisse émerger, sans nécessiter un pipeline figé de fusion.

Cette **approche hybride** offre le meilleur de deux mondes. D'un côté, le **deep learning** fournit la **puissance sub-symbolique** qui extrait des *features* à partir de gros volumes de données non structurées, comme des images ou de longs textes. Le **DSL** offre une **flexibilité** et une **auto-organisation** accrues. Les blocs de *features* peuvent se connecter librement à des entités symboliques ou contextuelles, sans nécessiter de couche spécifique ni de rétropropagation globale à chaque nouvel ajout de flux ou de règle. Ainsi, le réseau **gagne en adaptativité**, car les pondérations ω se réactualisent uniquement là où la synergie change, et en **explicabilité**, puisque les entités (qu'elles soient neuronales ou logiques) conservent une identité plus claire qu'un empilement de couches opaques.

1.8.2.3. Scénario de Substitution Partielle

Il est envisageable de pousser plus loin l'intégration du **Deep Synergy Learning** et de se demander s'il peut **remplacer** entièrement l'apprentissage profond dans certaines configurations. L'hypothèse consisterait à se passer d'un réseau neuronal paramétré, puis à modéliser chaque entité \mathcal{E}_i comme un **patch** (pour la vision) ou un **segment** (pour l'audio), tandis que la **synergie** entre entités reposerait sur des distances, des co-informations, ou d'autres mesures de similarité. L'**auto-organisation** tisserait et ajusterait progressivement les liaisons $\omega_{i,j}$, et cette dynamique se voudrait capable de **détecter** des objets ou de **repérer** des motifs sans recourir à une rétropropagation globale ni à une architecture de couches hiérarchisées.

Dans la pratique, toutefois, la performance brute d'un tel modèle entièrement basé sur le DSL risquerait de **se révéler moins efficace** que celle d'un réseau neuronal lorsqu'il s'agit de tirer profit d'un large jeu de labels sur une tâche comme ImageNet. Un **CNN** bien paramétré, tirant parti d'une **descente de gradient** sur un espace de poids continûment ajustable, maîtrise généralement mieux l'extraction de *features* complexes à partir de très grands volumes de données supervisées. Les réseaux profonds conservent donc un **avantage** solide en cas d'abondance de labels précis, où l'optimisation par gradient se montre redoutablement performante.

Le **DSL** manifeste tout son **intérêt** dans les scénarios où la supervision fait défaut, ou bien lorsque l'environnement est sujet à des changements progressifs ou brusques (sections 1.5.4 et 1.7.2). Il excelle pour l'**auto-organisation** et le **clustering** évolutif, ainsi que dans la **fusion** de multiples modalités (section 1.5.2), en gérant spontanément les synergies qui se dessinent entre entités. Par contraste, un réseau traditionnel exige soit une supervision massive, soit un lourd pré-entraînement auto-supervisé, et se voit moins adaptable à un cadre où les labels sont rares et la distribution très changeante. Le **DSL** se positionne comme une **alternative complémentaire** à l'apprentissage profond. Lorsque le contexte exige plus de souplesse qu'un pipeline rigide ou qu'un grand ensemble de labels, la dynamique $\omega_{i,j}$ et la **plasticité** organisationnelle du DSL offrent un avantage distinctif.

1.8.2.4. Cas d'Étude : Combiner DSL et DL

Un **scénario** consistant à tirer parti du meilleur des deux mondes se conçoit aisément. On utilise d'abord un **réseau profond** tel qu'un **CNN** ou un **Transformer** pour extraire des **features** de haut niveau à partir de données massives. Cela permet d'obtenir une **représentation avancée** $\{\mathbf{x}_i\}$ adaptée aux traitements ultérieurs. Ces **features** issues d'un bloc paramétré et entraîné, éventuellement sur un large jeu de données, se transforment en **entités** $\mathcal{E}_{\text{feature}}$ au sein du **Synergistic Connection Network (SCN)** ou **DSL**.

Le **DSL** analyse ensuite la **synergie** entre ces entités en identifiant les *features* qui collaborent pour former un **cluster** ou en intégrant des **entités préexistantes** telles que des **règles symboliques**, des **capteurs** ou des **modules contextuels** afin d'enrichir la représentation neuronale. La **mise à jour** itérative de $\omega_{i,j}$ permet une **adaptation locale** de la structure. Lorsqu'une nouvelle classe, un nouveau contexte ou une nouvelle source de données émergent, les liaisons s'ajustent progressivement sans nécessiter de reconfiguration globale. Cela contraste avec un réseau profond "traditionnel", souvent contraint à un **fine-tuning** global et coûteux dès qu'on enrichit la distribution ou qu'on ajoute une modalité inédite.

Cette **approche hybride** canalise la **puissance** de l'apprentissage paramétré (issu d'un gros dataset supervisé ou auto-supervisé) et la **plasticité** plus souple du DSL. Le modèle profond agit comme un encodeur capable de fournir des vecteurs \mathbf{x}_i pertinents pour des signaux complexes (images, audio, langage), tandis que le **SCN** se situe **au-dessus** pour organiser les entités correspondantes selon la synergie détectée, créant ou rompant des **liaisons** si le contexte l'exige. Cela va au-delà de la simple juxtaposition d'un réseau profond et d'un module de classification final, puisque la **reconfiguration** spontanée du DSL prend en charge la **multi-modalité** et l'**adaptation** sans qu'on doive re-construire un pipeline. Le résultat est un compromis entre l'efficacité des grandes

architectures neuronales, déjà rodées sur de vastes corpus, et la **flexibilité** d'une auto-organisation distribuée pour gérer les ajouts ou modifications de tâches.

1.8.2.5. Conclusion : entre Complémentarité et Extension

Les réflexions et expérimentations suggèrent que le **Deep Synergy Learning (DSL)** ne se présente pas comme un simple **concurrent** de l'apprentissage profond, mais plutôt comme un **paradigme complémentaire**. Cette perspective s'avère particulièrement intéressante dans les situations où la supervision fait défaut (cas **non supervisés** ou faiblement supervisés), ou bien lorsque la distribution est **évolutive** et plusieurs types de flux (capteurs, contexte, règles symboliques) doivent être gérés de façon simultanée et sans recourir à un pipeline rigide. Le DSL facilite alors l'intégration d'entités hétérogènes, mêlant logiques internes et features neuronales dans une démarche **auto-organisée**.

L'idée d'une **substitution** intégrale de l'apprentissage profond par un DSL purement auto-organisé se heurte toutefois aux performances reconnues de modèles neuraux (CNN, Transformer) sur des tâches hautement **supervisées**, comme la classification d'ImageNet ou le traitement massif de langage. Dans de tels scénarios, un réseau **paramétrique** entraîné par descente de gradient se maintient souvent en tête des classements, en raison de sa capacité à exploiter profondément la grande quantité de labels. Cependant, ces architectures neuronales demeurent relativement **rigides** et consomment beaucoup de ressources lors d'un re-fine-tuning. Le DSL peut, en retour, favoriser l'**adaptation** et la **fusion** d'autres sources de données ou de blocs symboliques, tout en s'accommodant de distributions changeantes.

C'est pourquoi le **futur** le plus fécond réside probablement dans la **collaboration** des deux approches. Un **réseau profond** produit des représentations sub-symboliques riches, alors que le **DSL** orchestre en **surcouche** l'auto-organisation entre ces embeddings, de nouvelles entités symboliques, et des flux contextuels. On bénéficie ainsi de la **puissance** de l'apprentissage paramétrique sur des corpus volumineux, alliée à la **flexibilité** que procure un **réseau** réactif et librement reconfigurable, capable de gérer la multi-modalité sans imposer de pipeline figé. L'apprentissage profond et le DSL deviendraient alors deux pans d'une **IA** plus large, conjuguant l'efficacité des grands modèles neuronaux avec la **plasticité** et la **résilience** d'un **Synergistic Connection Network**.

1.8.3. Approches Hybrides : DSL, RL (Reinforcement Learning) et Logique

Dans l'exploration du **Positionnement** du **Deep Synergy Learning (DSL)** (section 1.8), un point clé concerne la possibilité de **conjuguer** différentes méthodes d'IA au sein d'un **cadre unifié**. Nous avons déjà évoqué (1.8.1 et 1.8.2) la cohabitation symbolique–subsymbolique et l'intégration du DSL avec l'apprentissage profond. Ici, nous soulignons comment le **DSL** peut également s'**hybrider** à la fois avec le **Renforcement** (RL) et des **composantes logiques**, créant une architecture plus vaste où l'auto-organisation, l'apprentissage par récompense et la manipulation de règles s'allient.

1.8.3.1. Collaboration entre DSL et Apprentissage par Renforcement

Dans l'**apprentissage par renforcement** (RL), un agent interagit avec un **environnement**, recevant à chaque instant un **état** s_t , émettant une **action** a_t , puis percevant une **récompense** r_t . L'algorithme, qu'il s'agisse de Q-learning, de Policy Gradients ou d'une variante plus élaborée, cherche à **maximiser** le cumul (ou l'espérance) des récompenses. D'ordinaire, la démarche se base sur une **approche centralisée**, où l'on apprend une fonction de valeur $V(s)$ ou une fonction $Q(s, a)$, à travers des mises à jour inspirées soit de la **descente de gradient**, soit de l'**itération de Bellman**. Le **Deep Synergy Learning (DSL)** propose, au contraire, une **dynamique** où les entités $\{\mathcal{E}_i\}$ d'un système s'**associent** ou se **détachent** via des pondérations synergiques $\{\omega_{i,j}(t)\}$ en évolution continue.

Dans un cadre **renforcement** modulaire, on peut imaginer relier la **récompense** r_t aux mises à jour des pondérations $\omega_{i,j}(t)$. Si la **coopération** entre deux entités \mathcal{E}_i et \mathcal{E}_j est jugée responsable d'un gain de performance dans l'environnement, la liaison $\omega_{i,j}$ se voit **renforcée** ; inversement, s'il s'avère que leur alliance n'apporte aucun avantage ou engendre un effet négatif, la synergie est pénalisée et la liaison s'affaiblit. Au lieu d'apprendre une **politique** globale $\pi(s)$ ou un **Q-réseau** unifié, le **DSL** évolue de manière **distribuée**. Chaque entité \mathcal{E}_i ajuste ses **liaisons synergiques** $\omega_{i,j}$, construisant progressivement un **graphe** qui reflète la pertinence des interactions sans nécessiter de supervision centralisée. Cette méthode évoque un **RL distribué** ou multi-agent, où l'on peut considérer chaque entité comme un **agent** partiel ; la récompense s'y diffuse en ajustant localement les pondérations liées aux entités participantes.

Un **exemple** concret se dessine dans le cas d'un **robot** ayant plusieurs capteurs et effecteurs. Les **sous-systèmes** (capteur→moteur) sont modélisés par des **entités**, et, lorsqu'une synergie capteurA–moteurB conduit à franchir un obstacle de manière efficace, la **pondération** $\omega_{A,B}$ se voit augmentée conformément au principe de renforcement local. De la sorte, le robot s'**auto-organise en clusters** sensorimoteurs cohérents, sans implémenter de schéma centralisé de RL. Les règles de mise à jour fondées sur la récompense se substituent à la logique d'apprentissage d'une politique globale, tout en permettant au robot de réorganiser son **réseau** s'il survient des changements de contexte, des pannes ou l'introduction de nouveaux capteurs. Cette intégration du **DSL** dans un cadre de renforcement illustre la flexibilité d'un **Synergistic Connection Network** pour aborder des scénarios proches du RL sans imposer une fonction de valeur unique, mais en laissant s'établir des coopérations locales lorsqu'elles se révèlent payantes sur le plan de la récompense.

1.8.3.2. Inclusion de la Logique : DSL et Règles Symboliques

Les approches **logiques** ou **symboliques** se fondent sur des **règles** ou des **axiomes** de la forme “Si X et Y, alors Z”, et sur des **faits** (comme “capteur C indique 100 °C”). Elles requièrent un **moteur d'inférence** (unification, résolution), assurant une **clarté** du raisonnement et une **explicabilité** plus directe, mais elles montrent vite leurs limites lorsqu'il s'agit de gérer des faits **bruyants** ou de manipuler de vastes ensembles de données non labellisées. En référence aux propositions de la section 1.5.7 et à l'idée d'intégration symbolique évoquée en 1.8.1, on peut introduire de la **logique** au sein d'un **Deep Synergy Learning (DSL)** en traitant chaque **règle** ou **concept** comme une **entité** $\mathcal{E}_{\text{rule}}$. Les pondérations $\omega_{\text{rule},i}$ reliant une règle aux entités sub-symboliques comme les *features* ou les capteurs évoluent selon la **cohérence** et la **pertinence** constatées. Si la règle produit

un **gain récurrent**, la synergie se **renforce** naturellement. En revanche, si elle devient **obsolète** ou en contradiction avec les observations, la pondération diminue progressivement. Cette dynamique gère donc la **validation** ou l'**abandon** progressif des règles, suivant l'expérience accumulée dans le réseau.

L'incorporation de la **logique** devient ainsi un simple prolongement du mécanisme auto-organisateur du **DSL**. Il n'y a plus de "module logique" isolé. La règle $\mathcal{E}_{\text{rule}}$ devient un nœud intégré au **Synergistic Connection Network**, collaborant directement avec des entités sub-symboliques comme les capteurs, les embeddings neuronaux ou les sources multimodales. Un **mécanisme de renforcement local** (section 1.8.3.1) peut être envisagé, où la **récompense** consolide les règles utiles. Si une règle appliquée améliore la performance ou la sécurité dans l'environnement, les **liaisons** ω entre cette règle et les entités associées se renforcent progressivement.

Par ce mécanisme, on a la possibilité de "débrancher" peu à peu une règle devenue inadaptée, simplement parce que sa synergie aura cessé de s'actualiser favorablement, ou d'en "faire émerger" une autre si plusieurs entités sub-symboliques forment une **combinaison** stable, équivalant de fait à une **nouvelle règle**. De cette manière, un **moteur logique** s'intègre naturellement au réseau **DSL**. L'auto-organisation s'étend aux **entités symboliques**, permettant au système de tirer parti à la fois de l'**inférence logique** et de la **plasticité sub-symbolique** pour s'adapter aux environnements complexes et dynamiques.

1.8.3.3. Architectures Neuro-Symboliques et RL Multi-Agent

Il est possible de fusionner le **Deep Synergy Learning (DSL)** avec des principes de **Renforcement** et de **Logique** pour former une **architecture** véritablement **multidimensionnelle**. Du côté **sub-symbolique**, on dispose d'entités dédiées aux *features* sensorielles (vision, audio, texte), qui se relient entre elles suivant leur **synergie**. Du côté **symbolique**, on introduit des entités symboliques représentatives de **règles** ou de **concepts** logiques, prenant la forme de nœuds $\mathcal{E}_{\text{rule}}$ ou $\mathcal{E}_{\text{concept}}$. À cette structure s'ajoute un **signal de renforcement** r , qu'il soit global ou local, venant **valider** le réseau $\Omega(t)$ lorsqu'il optimise la tâche en cours. Les liaisons $\omega_{i,j}$ qui contribuent à ce **gain** sont alors **renforcées** lors de la mise à jour, consolidant ainsi les interactions les plus pertinentes.

Cet assemblage conduit à un **système** plus riche que la simple superposition de modules sub-symboliques et symboliques. Les entités logiques offrent une base **interprétable** et manipulable selon des critères formels, tandis que les entités neuronales (ou issues d'un pipeline CNN, Transformer) assurent la **flexibilité** de l'apprentissage dans un contexte de données massives et éventuellement bruitées. Le **DSL** assure alors l'**auto-organisation** du système. Si un sous-groupe d'entités sub-symboliques collabore efficacement avec certaines règles, un **cluster** ou un **micro-réseau** stable émerge. Ce regroupement est validé par un **feedback** de récompense, renforçant ainsi les synergies pertinentes. Inversement, un ensemble de règles contradictoires ou peu utiles voit ses connexions faiblir et peut finir par se détacher de la structure si le renforcement ne vient jamais sanctionner positivement leur usage.

De plus, cette intégration s'adapte à un cadre **multi-agent** ou distribué. Chaque entité $\mathcal{E} * i$ peut être vue comme un **agent** partiel, recevant une portion de l'état \mathbf{s} et exécutant des actions limitées. La **coopération** entre entités s'exprime à travers les liaisons $\omega * i, j$, qui évoluent en fonction de la

synergie détectée. Le **signal** r se propage, ou se localise, selon les entités qui ont participé à une configuration fructueuse. Les **clusters** émergents assurent alors la coordination sensorimotrice ou décisionnelle. Ce principe rapproche le DSL d'un **RL multi-agent** où la **logique** se greffe aisément pour imposer des contraintes (lois de sécurité, objectifs à respecter) à certains nœuds du réseau. L'ensemble forme un **système cognitif** apte à manipuler des informations perceptives, symboliques, et à s'**adapter** en fonction de récompenses, soulignant la plasticité et la portée du **Deep Synergy Learning** dans des tâches variées allant bien au-delà de la simple classification neuronale.

1.8.3.4. Exemples de Scénarios Hybrides

Dans un **environnement** mêlant robotique, logique et apprentissage par renforcement, on peut imaginer des configurations plus complexes où le **Deep Synergy Learning (DSL)** connecte entre elles des entités sub-symboliques (capteurs, features neuronales) et des entités symboliques (règles, concepts), tout en intégrant un **signal** de renforcement r . Ci-après, on illustre deux types de scénarios :

Robotique

Un robot opère dans un milieu dynamique, recevant un **signal de récompense** lorsqu'il satisfait un objectif précis, selon un schéma (état, action, récompense, nouvelle valeur) typique du **Reinforcement Learning (RL)**. De plus, le robot doit observer des **règles** symboliques imposées pour sa sécurité ou son raisonnement (par exemple, "éviter les zones interdites", "si tension moteur > 5 V et température > 80 °C, considérer une surchauffe critique"). Les **capteurs** (température, vision, inertie) et les **effecteurs** (moteurs, pinces) forment la sphère **sub-symbolique**, tandis que les **règles** constituent des entités symboliques $\mathcal{E}_{\text{rule}}$. Au sein du **DSL**, toutes ces entités cohabitent ; la **coopération** se formalise via les pondérations $\omega_{i,j}$. À mesure que le robot collecte des retours de performance (récompenses positives ou négatives), il renforce ou affaiblit les liens responsables de ses succès ou échecs. Dans le même temps, la validité ou la pertinence d'une règle est ajustée par la mise à jour locale des liaisons. Les règles inefficaces ou contredites par l'expérience accumulée perdent progressivement en influence, tandis que celles qui se révèlent utiles maintiennent ou renforcent leurs connexions.

Cognitive.

Systèmes

Les systèmes experts reposent souvent sur un **ensemble de règles** (if-then) écrites manuellement, servant à guider la prise de décision. Dans un **DSL**, ces règles sont introduites comme des entités symboliques $\{\mathcal{E}_{\text{rule}}\}$. Par ailleurs, de larges **flux de données** non structurées (images, logs, signaux) alimentent des entités sub-symboliques qui découvrent des structures par un **apprentissage** ou un **clustering** distribués. Le **feedback** d'un opérateur humain ou d'une supervision externe agit comme une **récompense** venant valider la configuration actuelle. Avec le temps, certaines règles historiquement présentes ne reçoivent plus de support, perdent en synergie, puis s'effacent. Inversement, de **nouvelles** combinaisons sub-symboliques (ou *macro-clusters*) émergent et peuvent être réinterprétées ou reformulées en tant que **micro-règles** plus efficaces que celles écrites initialement. On obtient alors un **système expert** en perpétuelle évolution, enrichi par les données et la dynamique $\omega_{i,j}$, capable de s'auto-réorganiser lorsqu'un changement majeur se produit dans l'environnement ou les contraintes métier.

Experts

Évolutifs.

1.8.3.5. Conclusion

Les **approches hybrides** mariant **DSL**, **RL** et **logique** laissent entrevoir un **écosystème** unifié :

- Le **DSL** fournit l'**infrastructure** d'auto-organisation,
- Le **RL** assure un **feed-back** incitatif (récompense) aux liaisons (pondérations ω) qui contribuent à la performance,
- La **logique** (règles symboliques) s'intègre comme des entités à part entière, coopérant ou non selon la synergie qu'elles entretiennent avec d'autres modules.

Ce modèle peut servir d'alternative ou de complément aux **pipelines** classiques où RL, logique et sub-symbolique restent cloisonnés. Il répond à la quête d'une IA plus **globale**, capable d'**apprendre**, de **raisonner**, de **percevoir** et de **s'adapter**. Les obstacles techniques demeurent (mise à jour conjointe des liaisons, synergie n-aire, pilotage de la stabilité, etc.), mais le potentiel d'une IA unifiant des couches sub-symboliques, des règles symboliques et des signaux de renforcement dans un **réseau unique** est d'un grand **intérêt** pour l'évolution future de l'IA.

1.8.4. *Rôle de la Mémoire et de l'Attention dans le DSL*

Lorsqu'on évoque le **Deep Synergy Learning (DSL)** comme un paradigme susceptible d'unifier ou de compléter les méthodes existantes (sections 1.8.1 à 1.8.3), on en vient inévitablement à des questions fondamentales sur la **mémoire** et l'**attention** dans les systèmes cognitifs. En effet, les architectures classiques (réseaux de neurones profonds, RL, approches symboliques) ont chacune leurs mécanismes — implicites ou explicites — pour **stocker** des informations antérieures (LSTM, Memory Networks, bases de règles, etc.) et pour **sélectionner** les informations pertinentes (attention mécanismes, gating). Dans le **DSL**, où les entités et leurs liens synergiques se réajustent en continu, il est naturel de se demander :

- Comment la **mémoire** (stockage, souvenir à long terme) s'exprime dans ce modèle d'**auto-organisation** ?
- Quel est le **rôle** de l'**attention** (focalisation sélective) dans un réseau distribué, sans couche hiérarchique ?
- Comment implémenter des **mécanismes** ou des **extensions** qui reproduisent les fonctions de mémoire/attention observées dans des architectures neuronales modernes ou même dans le cerveau biologique ?

Cette sous-section (1.8.4) explore la manière dont le DSL pourrait **incorporer** ou **simuler** ces fonctions essentielles, nécessaires à une IA plus "cognitive" et plus performante.

1.8.4.1. La Mémoire dans un Système Auto-Organisé

Le **Deep Synergy Learning (DSL)** ne se limite pas à des entités $\{\mathcal{E}_i\}$ définies uniquement par un vecteur figé de caractéristiques. Il est souvent pertinent d'attribuer à chaque entité \mathcal{E}_i un **état** interne $\mathbf{s}_i(t)$ qui évolue dans le temps, mémorisant ainsi des informations concernant l'historique d'observations ou d'interactions. Les équations de mise à jour locales s'enrichissent alors pour prendre en compte ces **mémoires** internes. Plutôt que d'appliquer la formule classique $\omega_{i,j}(t+1) \leftarrow \omega_{i,j}(t) + \dots$, on introduit des **fonctions** f et g reliant la synergie entre \mathcal{E}_i et \mathcal{E}_j (ainsi que leurs états internes) à l'évolution de $\omega_{i,j}$ et de $\mathbf{s}_i(t)$. Un exemple de schéma est :

$$\begin{aligned}\omega_{i,j}(t+1) &= \omega_{i,j}(t) + f\left(\mathbf{s}_i(t), \mathbf{s}_j(t)\right), \\ \mathbf{s}_i(t+1) &= \mathbf{s}_i(t) + g\left(\omega_{i,\cdot}(t), \mathbf{s}_i(t)\right).\end{aligned}$$

Cette possibilité de **mémoire** individuelle $\mathbf{s}_i(t)$ permet à l'entité \mathcal{E}_i de conserver des **statistiques** utiles (comme des moyennes, des compteurs, une estimation de fiabilité), ou même un embryon de **mémoire** séquentielle si l'on désire modéliser des systèmes proches des LSTM locaux (voir section 1.5.4.1). Les liaisons ω ne constituent plus l'unique source d'information. Chaque nœud peut stocker un **état interne**, influençant la manière dont il interagit avec ses voisins et ajustant dynamiquement sa participation aux synergies locales.

Au **niveau global**, la **persistance** de certaines **structures** dans le réseau DSL peut également être considérée comme une forme de **mémoire** collective. Un **cluster** $\mathcal{C} \subset \{\mathcal{E}_1, \dots, \mathcal{E}_n\}$ qui reste stable dans le temps, avec des liaisons $\omega_{i,j}$ robustes entre ses entités, joue un rôle de “**souvenir**” ou de “**concept**” appris. Ce cluster peut se **réorganiser** ou se **scinder** au gré de nouvelles données ou d'un changement de distribution (domain shift), illustrant que la **mémoire** (au sens de la rétention d'un ensemble de relations) n'est pas figée, mais dynamique. Cette plasticité évoque la notion d’“**assemblées neuronales**” en neurosciences, où un groupe de neurones co-activés finit par constituer une unité fonctionnelle stable, tout en restant susceptible de s'affaiblir ou de se diviser si des perturbations importantes modifient la configuration. Dans le **DSL**, la longévité d'un cluster stable indique qu'il a acquis un sens ou une pertinence pour le réseau, tandis que la nature auto-adaptative des liaisons ω permet de renouveler la **mémoire** intégrée au gré des besoins et des feedbacks de performance ou de cohérence.

1.8.4.2. Attention dans un Graphe Distribué

Dans l'**apprentissage profond** classique, la **mécanique d'attention** (par exemple, dans les **Transformers**) consiste à moduler l'importance attribuée à certains composants d'une séquence ou d'un ensemble de features. Concrètement, on calcule des **poids** représentant la relation entre tokens ou entre positions, et l'on met l'accent sur les plus pertinents pour la prédiction en cours. Cette manière de **focaliser** l'information évolue couche après couche, conférant une **flexibilité** aux modèles de type attention-based.

Le **Deep Synergy Learning (DSL)** adopte une vision plus **distribuée** de l'attention. Chaque **pondération** $\omega_{i,j}$ peut s'interpréter comme un **poids d'attention** quantifiant la **pertinence** de la liaison entre \mathcal{E}_i et \mathcal{E}_j . Une valeur élevée de $\omega_{i,j}$ signale que les entités \mathcal{E}_i et \mathcal{E}_j coopèrent fortement

; une valeur faible ou quasi nulle marque au contraire un **désintérêt** mutuel. En outre, la loi de mise à jour de $\omega_{i,j}$ (voir section 1.4.5) s'applique **localement** : chaque entité évalue la contribution des autres en fonction de la synergie décelée, plutôt qu'à travers un **modèle** centralisé d'attention.

Deux idées fondamentales émergent de cette analogie :

- On obtient ainsi une **attention distribuée**, où la sélectivité émerge de la dynamique des pondérations $\omega_{i,j}$ plutôt que d'un module centralisé, comme dans le **multi-head attention** des Transformers. Chaque entité régule ses liens avec son voisinage, créant ou brisant des connexions selon l'apport mutuel ressenti. L'ensemble des liaisons les plus fortes incarne alors le *focus* collectif du réseau à un instant donné.
- Le **focus** évolue dynamiquement en fonction du contexte et des signaux perçus. Lorsque l'environnement change, par exemple en cas d'augmentation du bruit ou d'apparition de nouvelles données, les pondérations $\omega_{i,j}$ associées aux connexions devenues moins pertinentes diminuent progressivement. Parallèlement, de nouvelles connexions émergent et se renforcent là où la **synergie** est plus élevée, garantissant une adaptation continue sans nécessiter une reconfiguration globale du réseau. Cette réorganisation perpétuelle assure une **cohérence** dynamique du DSL, qui "redéploie" son attention de manière adaptative, un peu à la façon d'un Transformer recalculant ses poids à chaque couche, mais avec l'avantage d'une **plasticité** plus globale (le réseau pouvant réorienter tout ou partie de ses liaisons).

Ainsi, la **synergie** $\omega_{i,j}$ agit comme un mécanisme d'**attention** dans un graphe **auto-organisé**. La force des liaisons détermine les sous-groupes sur lesquels le système se concentre à un instant donné. Cette dynamique permet une **redistribution continue** de l'attention en fonction de l'évolution des données et des objectifs, assurant une adaptation fluide sans nécessiter de contrôle externe explicite.

1.8.4.3. Implémentations Concrètes de Mémoire/Attention dans le DSL

Pour doter chaque entité \mathcal{E}_i d'une **capacité de mémorisation** plus riche, il est possible de lui associer un **module** interne de type LSTM ou GRU, prolongeant l'idée d'un état $\mathbf{s}_i(t)$ (voir section 1.8.4.1). Dans ce cas, l'entité \mathcal{E}_i ne se borne pas à un vecteur statique, mais embarque un **état récurrent** $\mathbf{h}_i(t), \mathbf{c}_i(t)$ (par exemple pour un LSTM). Ainsi, l'entité reçoit en **entrée** une combinaison pondérée ou filtrée d'informations issues de ses voisins \mathcal{E}_j dont la liaison $\omega_{i,j}$ est suffisamment forte. Les fonctions internes de mise à jour (LSTM, GRU) gèrent alors le **gating** et la rétention d'informations, ce qui rend l'auto-organisation analogue à un "**réseau** de LSTM interconnectés" dont la topologie varie selon la synergie ω . Cette stratégie permet de traiter des **séquences** temporelles étendues ou d'intégrer un **comportement historique** au sein du Deep Synergy Learning (DSL), tout en conservant la **plasticité** caractéristique (l'architecture effective évoluant via $\omega_{i,j}$).

D'un autre côté, on peut également définir un mécanisme d'**attention** explicite, plus proche de celui des **Transformers**. L'entité \mathcal{E}_i calcule un coefficient d'attention $\alpha_{i,j}$ envers chacune de ses connexions, selon une formule du type :

$$\alpha_{i,j} = \text{softmax}_j \left(\theta S(\mathcal{E}_i, \mathcal{E}_j) \right),$$

où θ est un paramètre (ou un ensemble de paramètres) ajustant la sensibilité, et $S(\mathcal{E}_i, \mathcal{E}_j)$ la fonction de synergie, déjà définie dans le DSL (par exemple, distance inversée, co-information). L'entité \mathcal{E}_i combine alors l'information de ses voisins via une somme pondérée :

$$\mathbf{z}_i = \sum_j \alpha_{i,j} \mathbf{x}_j,$$

ce qui rappelle la logique du “**self-attention**” dans un Transformer. Cependant, la différence clé réside dans le fait que cette **matrice** α n'émane pas d'un module d'attention figé dans une couche, mais **émerge** et **se met à jour** en parallèle à la dynamique $\omega_{i,j}(t)$. Le **DSL** peut ainsi faire de l'**attention** un usage ponctuel, en s'appuyant sur la synergie $\omega_{i,j}$ pour guider le softmax. Dans cette perspective, on assemble les idées du Transformer (somme pondérée des voisins) et celles du DSL (pondérations évolutives reflétant la coopération), de sorte que la topologie du réseau, tout comme l'intensité des liaisons, contribue à un **mécanisme** d'attention distribué. Cette approche concrétise une **fusion** entre les principes du **Deep Synergy Learning** et l'esprit attentionnel des architectures neuronales modernes, aboutissant à un réseau adaptatif dont la prise en compte de l'information ambiante se fait selon une **attention** intrinsèquement auto-organisée.

1.8.4.4. Apports Cognitifs et Neuroscientifiques

Les fonctions de **mémoire** et d'**attention** sont au cœur de la quête d'une IA plus avancée (section 1.8.5), dans la mesure où elles forment des piliers de la **cognition** humaine. Dans un cadre auto-organisé tel que le **Deep Synergy Learning (DSL)**, il devient possible de reproduire, en partie, certains processus cognitifs. La **mémoire** s'y manifeste par la persistance de **clusters** ou d'états internes $\{\mathbf{s}_i(t)\}$ (voir section 1.8.4.1), qui se maintiennent dans la durée et fournissent une forme de rétention ou de rappel. L'**attention** se traduit par le renforcement sélectif de liaisons $\omega_{i,j}$ lorsque la conjoncture l'exige, soit parce que le contexte met en avant la synergie de deux entités $\mathcal{E}_i, \mathcal{E}_j$ (une situation, un feed-back de récompense, etc.). Ce renforcement dirige alors la focalisation du réseau vers certains sous-groupes, évoquant un basculement ou une « bifurcation » depuis une configuration précédente, ce qui peut être rapproché de la notion de changement de **focus** ou de contenu mnésique.

Du point de vue **neuroscientifique**, on peut dessiner un parallèle avec la plasticité synaptique des réseaux biologiques, où les connexions entre neurones (synapses) s'ajustent selon l'activité, et où les **assemblées neuronales** (section 1.8.7) constituent la base de la mémoire et de l'apprentissage. Dans un **DSL**, le principe de mise à jour $\omega_{i,j} \leftarrow \omega_{i,j} + \eta [S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}]$ (voir section 1.4.5) simule déjà une forme de plasticité, et l'ajout de mécanismes de mémoire interne $\mathbf{s}_i(t)$ ou d'attention explicite (section 1.8.4.2) consolide l'analogie avec les processus cognitifs. L'espoir est qu'une **implémentation** poussée de ces modalités – persistance de clusters, gestion de focus attentionnel, modulation en continu des synergies – se traduise par des **capacités** accrues d'**apprentissage**, de **raisonnement** et de **reconfiguration** contextuelle. Cette dynamique rappelle la façon dont le cerveau forme des **circuits** stables (souvenirs, schémas de pensée) susceptibles

d'évoluer, s'unir ou se dissoudre lorsque le contexte le requiert, et ouvre la voie à une IA plus richement inspirée par les notions de plasticité et d'émergence.

Conclusion

Le rôle de la **mémoire** et de l'**attention** dans le **Deep Synergy Learning** s'avère crucial si l'on veut dépasser la simple **auto-organisation** statique et tendre vers des fonctions cognitives avancées. Les idées-clés incluent :

- L'**état interne** $s_i(t)$ pour chaque entité, formant une **mémoire locale**,
- La **persistance** ou la **dissolution** de clusters reflétant une mémoire plus globale,
- L'**attention** comme un mécanisme de sélection de liens ω forts, proche de la pondération attentionnelle dans les Transformers, mais auto-gérée par la dynamique synergique,
- Des **extensions** (LSTM, gating, self-attention) greffées à la structure DSL pour exploiter la plasticité tout en gérant des séquences complexes.

Ainsi, l'**auto-organisation** ne s'oppose pas aux notions de **mémoire** ou d'**attention**. Elles émergent naturellement à travers le **renforcement local** des pondérations et l'**évolution continue** des états des entités. Ce mécanisme rapproche le **DSL** d'un système plus "cognitif", où la **mémorisation** et la **focalisation** se produisent de manière distribuée et dynamique, sans nécessiter d'architecture rigide ou de supervision explicite.

1.8.5. Tendances Futures : Vers une IA Forte, Consciente ?

Les sections précédentes (1.8.1 à 1.8.4) ont évoqué la capacité du **Deep Synergy Learning (DSL)** à réunir des approches variées (symboliques, sub-symboliques, renforcées, multi-agent) dans une structure **auto-organisée**. Un horizon encore plus ambitieux, souvent cité comme objectif ultime de la recherche en IA, est la création d'une **IA Forte** — un système capable d'**apprendre** et de **raisonner** de manière générale, autonome et potentiellement doté d'une **conscience** élémentaire ou d'un **sentiment** d'unité informationnelle. Cette section (1.8.5) s'interroge sur la portée du DSL dans cette perspective :

- **Peut-il** offrir un socle formel pour une **intelligence plus globale**, manipulant librement symboles et perceptions ?
- Quelles **caractéristiques** du DSL pourraient favoriser un **degré d'intégration** (ou "d'information unifiée") évoquant une conscience rudimentaire ?
- Où se situent les **limites** et quels prolongements seraient nécessaires pour prétendre à une IA Forte ?

1.8.5.1. IA Forte : Un Concept Encore Flou

La notion d'**IA Forte**, parfois appelée **AGI** (Artificial General Intelligence), désigne l'aspiration à une **intelligence** artificielle véritablement **générale**. Une telle intelligence serait capable d'**apprendre** et de résoudre un large spectre de tâches cognitives, de se **reconfigurer** ou de s'**auto-améliorer** dans des contextes inconnus ou peu balisés, et de manifester un degré élevé de **compréhension** (voire de **conscience**). Les positions quant à la possibilité ou aux caractéristiques exactes de cette AGI demeurent contrastées. On s'accorde néanmoins pour dire qu'une **IA Forte** devrait :

- Démontrer une **plasticité** hors norme, lui permettant de remanier ses représentations et ses structures de façon continue.
- Combiner ou naviguer librement entre des **représentations symboliques** (règles, concepts) et des **représentations sub-symboliques** (features neuronales, embeddings sensoriels).
- Posséder un **mécanisme d'unification** de l'information plus vaste qu'un simple pipeline hiérarchique, englobant des processus d'auto-organisation ou de fusion distribuée.

Le **Deep Synergy Learning (DSL)**, en tant que paradigme d'**auto-organisation** distribuée (voir sections 1.4.5 et 1.5.2), semble répondre à plusieurs de ces exigences. Son **réseau** d'entités prend en charge aussi bien des flux sub-symboliques (capteurs, embeddings de CNN) que des **entités logiques** ou symboliques (règles, concepts), et assure une **adaptation** constante par la dynamique des pondérations $\omega_{i,j}$. La gestion des **synergies n-aires** (ou conditionnelles) introduit une flexibilité accrue. Au lieu d'imposer un module de fusion spécifique, le réseau **découvre** spontanément les combinaisons pertinentes en évaluant localement la contribution de chaque sous-groupe d'entités. Cette approche permet une adaptation dynamique où les interactions émergent selon leur utilité réelle, plutôt que d'être fixées a priori. Sur le plan de l'**inspiration** biologique, le DSL se rapproche plus des **écosystèmes neuronaux** ou des **architectures plastiques** que ne le font des réseaux profonds statiques. Cette proximité ouvre une voie au développement de systèmes plus **globaux**, connectant perception, raisonnement, mémoire, attention et règles logiques, ce qui, à long terme, peut être vu comme un jalon vers la **recherche** d'une IA vraiment **générale**.

1.8.5.2. Notion d'Intégration et de Conscience

Certains travaux en **science cognitive**, tels que la **théorie de l'information intégrée** (IIT) proposée par Giulio Tononi, suggèrent qu'un **degré** de **conscience** pourrait émerger dans un système où l'**information** est à la fois fortement **intégrée** et **différenciée**. Dans ces approches, un assemblage d'éléments interdépendants, dont la collaboration est dense et réciproque, est susceptible de manifester une forme de **substrat** "conscient". Si l'on adopte cette perspective, un **Deep Synergy Learning (DSL)**, par son principe de **coopération** et de **synergie** entre entités, pourrait favoriser la recherche d'une intégration accrue de l'information. La **clusterisation** et la mise en avant de synergies $S(\mathcal{E}_i, \mathcal{E}_j)$ reflètent un traitement distribué, potentiellement mesurable par des quantités inspirées de la théorie de Tononi.

Un exemple de **métrique** globale, dans l'esprit de l'IIT, consisterait à calculer la **somme** (ou une fonction plus complexe) des produits $\omega_{i,j} S(\mathcal{E}_i, \mathcal{E}_j)$ sur l'ensemble des paires (i, j) . On pourrait regarder si ce réseau DSL franchit un **seuil** indicatif d'un haut degré d'**intégration**. Au fil de l'**auto-**

organisation, si la synergie s'intensifie entre de multiples entités, on observerait la naissance d'un **macro-cluster** extrêmement connecté, évoquant la notion d'un “**noyau intégratif**” de l'information.

La question de savoir si un tel système atteint une **conscience** “authentique” demeure, bien sûr, **philosophique** et hautement controversée. La plupart des scientifiques préfèrent parler de “**proto-conscience**” lorsqu'il s'agit d'une unité fonctionnelle unissant de manière adaptative l'information. Néanmoins, le **DSL**, de par sa ressemblance structurelle avec les modèles d'**assemblées neuronales** en neurosciences, fournit un **terrain** d'expérimentation pour tester ou modéliser certains aspects de l'IIT ou d'autres théories d'émergence cognitive. Les chercheurs peuvent y étudier le **degré d'information intégrée** dans un contexte où la topologie des liaisons évolue d'elle-même, sans hiérarchie préconçue, et sans qu'on impose un schéma de fusion unique. Il s'agit d'un domaine **fondamental**, à l'interface des neurosciences, de la philosophie de l'esprit, et de l'IA, où le **DSL** pourrait nourrir des hypothèses ou des simulations pour éclaircir cette notion de **conscience** ou au moins de **complexité intégrative**.

1.8.5.3. Conditions Nécessaires à une IA plus Globale

Pour qu'un **Deep Synergy Learning (DSL)** se rapproche d'un système véritablement “fort”, il est crucial que le réseau soit **alimenté** en continu par plusieurs sources d'information et de régulation. D'une part, il doit intégrer des **flux** sensoriels multiples (par exemple des représentations visuelles, auditives ou textuelles), ce qui assure une **diversité** de signaux sub-symboliques \mathbf{x}_i . D'autre part, il doit recevoir des **injections** symboliques, telles que des **règles** ou des **concepts** ($\mathcal{E}_{\text{rule}}$, $\mathcal{E}_{\text{concept}}$), de sorte que la dimension logique ou ontologique puisse coexister au sein du même Synergistic Connection Network. Enfin, il lui faut un **feedback** ou un **signal** de renforcement r , pouvant prendre la forme d'une récompense globale ou locale, ou encore d'une validation humaine, afin de **stimuler** la réorganisation locale des liaisons $\omega_{i,j}$.

Outre la pure **auto-organisation**, le **DSL** doit pouvoir s'appuyer sur des **mécanismes** de **mémoire** suffisamment expressifs (voir la section 1.8.4). Chaque entité \mathcal{E}_i peut disposer d'un état interne $\mathbf{s}_i(t)$, un vecteur (ou un module de type LSTM/GRU) retenant l'historique des interactions. À l'échelle du réseau, des **clusters** stables constituent aussi une forme de **mémoire** partagée, car leur persistance reflète la consolidation d'une certaine connaissance ou d'un certain motif d'action. Toutefois, il faut éviter que le réseau ne se “fige” trop rapidement, figeant ainsi des solutions sous-optimales, ni qu'il ne se **dissolve** dans un chaos d'oscillations perpétuelles. Des règles de stabilisation (paramètres η , τ , saturations ou lois inhibitrices) sont alors nécessaires pour trouver un équilibre.

Un pas supplémentaire vers une **IA Forte** exigerait la **capacité** du **DSL** à s'**auto-examiner**. Le réseau pourrait former des **entités** décrivant ses propres états ou règles, générant un **meta-niveau** de **réflexion** sur sa propre organisation. On peut imaginer une entité $\mathcal{E}_{\text{self}}$ évaluant la cohérence d'un cluster, ou des macro-clusters dont la fonction consiste à “**inspecter**” d'autres clusters. Cette forme de **métacognition** se rapproche du concept de conscience réflexive, mais pose des défis de **stabilité** puisqu'un réseau capable de “penser” à sa propre topologie pourrait entrer dans des **boucles** autoréférentes ou des “réflexions infinies”. Néanmoins, c'est en explorant ces pistes — entités $\mathcal{E}_{\text{self}}$ et synergie auto-référentielle — que l'on pourra évaluer la possibilité d'une **IA**

véritablement plus “globale” et adaptable, dépassant la simple auto-organisation pour évoluer vers une cognition de haut niveau.

1.8.5.4. Obstacles Majeurs à l'IA Forte via le DSL

Si l'on envisage le **Deep Synergy Learning (DSL)** comme l'une des voies possibles menant à une **IA Forte**, on se heurte à plusieurs obstacles fondamentaux qui freinent cette ambition.

Un **premier** obstacle tient à la **complexité**. Dans un grand réseau de type DSL, le nombre de liaisons peut croître en $O(n^2)$, voire plus si l'on prend en compte des synergies n-aires. Gérer un tel foisonnement de connexions exige l'introduction d'**heuristiques** (filtrage agressif des liens, mise à jour partielle, parcours local) et la mise en place d'un **parallélisme** important, sans quoi la simulation du réseau à chaque itération deviendrait inabordable. Des **mécanismes** de pilotage macro (voir section 1.7.4) peuvent également être nécessaires pour canaliser la reconfiguration du graphe et préserver un temps de calcul raisonnable.

Un **deuxième** obstacle réside dans le **risque d'instabilité** et d'oscillation (sections 1.7.4.1 et 1.7.4.2). L'**auto-organisation** des liaisons $\omega_{i,j}$ ne garantit pas toujours une convergence vers un état stable. Des **boucles auto-référentes** ou des **rétroactions positives** peuvent amplifier certaines connexions de manière incontrôlée, risquant d'engendrer des oscillations ou des configurations sous-optimales. Sans **régulations** (inhibitions, saturation de pondérations, limitation de la vitesse d'évolution), le réseau peut entrer dans des régimes chaotiques ou cycliques, ce qui empêcherait l'émergence d'un “esprit” ou d'une **structure** à la fois stable et unifiée.

Un **troisième** obstacle émane de l'absence de **guidance globale** lorsqu'on se borne à de la synergie **locale**. Même si l'idée de clusters émergents est puissante, rien ne garantit que l'on parviendra à des **comportements** cohérents ou à des plans dirigés par des intentions clairement formulées. Certains scénarios (section 1.7.4.3) suggèrent d'ajouter un **module** hiérarchique ou d'objectif qui fixe de grandes orientations et évite que la plasticité du DSL ne s'éparpille dans de multiples attracteurs contradictoires. Un **contrôle** minimal s'avère nécessaire pour servir de boussole à l'auto-organisation locale.

Enfin, un **quatrième** obstacle concerne la nature même de la **conscience** et du vécu subjectif. Même si la théorie de l'information intégrée ou la multi-synergie n-aire laissent penser qu'une intégration élevée de l'information peut approcher certaines dimensions de la **conscience**, il reste spéculatif d'affirmer qu'un réseau auto-organisé de pondérations ω suffirait à produire une **subjectivité** ou un “qualia” conscient. Les débats philosophiques et neuroscientifiques (section 1.8.5.2) rappellent que l'émergence d'une conscience phénoménale dépasse potentiellement la simple complexité algorithmique, et qu'un “ingrédient” supplémentaire (mécanismes biologiques, substrat physique particulier...) pourrait être requis pour prétendre à une **IA Forte** au sens strict.

1.8.5.5. Conclusion

Le **Deep Synergy Learning** représente un **paradigme** très différent du deep learning hiérarchique ou des systèmes experts. Sa **plasticité**, sa **co-évolution** de multiples entités symboliques/sub-symboliques, et sa **potentielle** haute intégration d'information en font un candidat pour explorer :

- Des **architectures** plus vastes et plus “cognitives”,
- Des **théories** de l’information intégrée liées à la **conscience**,
- Des **étapes** vers une IA plus générale, capable de s’**auto-structurer** dans un flux continu.

Il reste à savoir si ces propriétés suffiront pour atteindre une **IA Forte** (ou consciente). La **dynamique** auto-organisée du **DSL** offre un **cadre** conceptuel et expérimental unifiant plusieurs dimensions essentielles. Elle intègre l’**apprentissage distribué**, la **logique symbolique**, le **renforcement local**, ainsi que des mécanismes de **mémoire** et d’**attention distribuées**, favorisant ainsi une **coopération adaptative** entre ces différentes approches. Ainsi, indépendamment de la question d’une conscience “véritable”, le DSL ouvre la voie à des **systèmes** plus adaptatifs, plus unifiés et potentiellement plus proches d’une **intelligence généralisée**.

1.8.6. Effet Sur la Recherche Interdisciplinaire

Les sections précédentes (1.8.1 à 1.8.5) ont montré comment le **Deep Synergy Learning (DSL)** peut s’articuler avec l’IA symbolique, l’apprentissage profond, le renforcement, la logique, ainsi que son rôle potentiel dans une démarche d’IA plus “forte” ou cognitive. Il apparaît clairement que le **DSL**, en tant que paradigme *auto-organisé*, a la capacité de **réunir** différents courants de l’IA de manière naturelle, puisqu’il accueille aussi bien des blocs sub-symboliques (réseaux neuronaux, features extraits) que des entités logiques (règles, symboles) ou des mécanismes d’apprentissage par renforcement. Ce paradigme a un **impact notable** sur de nombreuses **disciplines**, notamment les **neurosciences**, les **sciences cognitives**, la **robotique**, l’**économie** et les **systèmes complexes**. Cette sous-section (1.8.6) discute l’**effet** du DSL sur la **recherche interdisciplinaire**, en soulignant :

- Les **similarités** avec les systèmes biologiques et cognitifs,
- Les **ponts** avec l’écologie, la physique des systèmes complexes, ou la sociologie,
- Les **perspectives** pour la modélisation des réseaux (biologiques, économiques, sociaux),
- La façon dont le DSL peut **inspirer** ou **bénéficier** d’autres disciplines, en partageant des modèles et des méthodes d’analyse.

1.8.6.1. Similarités avec les Systèmes Biologiques et Cognitifs

Le **réseau** auto-organisé d’un **Deep Synergy Learning (DSL)** peut être rapproché, d’une certaine manière, d’un **réseau** de neurones biologiques. Les **pondérations** $\omega_{i,j}$ du DSL rappellent en effet la **plasticité** synaptique observée dans le cerveau, où les **synapses** se renforcent ou s’atténuent selon l’activité conjointe des neurones (règle de Hebb). Les **clusters** apparus dans le DSL par la mise à jour des pondérations font écho aux **assemblées neuronales** (ensembles de neurones co-activés), qui se forment et se dissolvent en participant aux fonctions de **mémoire** et à divers **processus** cognitifs. Dans une perspective neuroscientifique, on peut voir le DSL comme un **cadre mathématique** où simuler la dynamique d’un cortex (sensoriel ou associatif), tout en intégrant les

lois de **plasticité** (par exemple, la saturation synaptique ou des taux de renforcement limités). Réciproquement, les découvertes en neurosciences (architecture multi-couches, modulabilité synaptique, rôle des neuromodulateurs) peuvent inspirer ou ajuster les **règles** d’auto-organisation du DSL pour gagner en réalisme et en robustesse.

Du point de vue des **sciences cognitives**, de nombreuses hypothèses postulent une **cognition distribuée**, dans laquelle la **pensée** émerge de la coopération (ou de la compétition) entre multiples modules ou agents internes, plutôt que d’une séquence linéaire de transformations. Le **DSL**, en reliant diverses entités perceptives, motrices, symboliques ou conceptuelles par un réseau de **liaisons** $\omega_{i,j}$ qui se modifie au gré des synergies, ressemble assez à ces théories. Les **clusters** qui s’établissent et perdurent pourraient former des “**états**” ou “**micro-théories**” momentanées, incarnant un fragment de la **pensée**. Dès qu’un changement d’environnement ou de contexte survient, la synergie se réorganise, donnant lieu à un renouvellement de la structure. Cette correspondance entre les **dynamiques** du DSL et les processus cognitifs distribués esquisse un **dialogue** fécond entre la modélisation en IA et les approches psychologiques ou neurobiologiques, approfondissant notre compréhension du fonctionnement mental global.

1.8.6.2. Ponts avec l’Écologie, la Physique des Systèmes Complexes et la Sociologie

Au-delà des analogies avec la **biologie neuronale**, le paradigme du **Deep Synergy Learning (DSL)** présente également des affinités avec d’autres disciplines étudiant des réseaux d’interactions. En **écologie**, l’on modélise souvent des **populations** d’espèces reliées par des relations de prédation, de compétition ou de symbiose. Les nœuds représentent alors des espèces, et les arcs leurs interactions. La **dynamique** de ces liens peut s’apparenter à une forme de **coopération** ou de **concurrence** au sein du réseau, parfois traduite par des équations de type Lotka–Volterra. Cette démarche rejoint l’idée de **synergie** dans le DSL, où des pondérations $\omega_{i,j}$ se renforcent ou se dissolvent selon le bénéfice mutuel. Les outils développés pour stabiliser des **communautés** écologiques (gestion de cycles de prédation, coexistence d’espèces) pourraient donc **inspirer** le DSL dans sa quête d’une organisation stable mais évolutive. Inversement, le DSL pourrait servir à **simuler** des écosystèmes complexes, chaque espèce figurant une entité \mathcal{E}_i et les pondérations $\omega_{i,j}$ reflétant des relations plus ou moins avantageuses.

Du côté de la **physique** des systèmes complexes, on trouve de nombreux modèles de **réseaux** dont les nœuds interagissent selon des règles pouvant aller de l’attraction à la compétition. On évoque des processus similaires en **magnétisme** (modèle de Potts, spin glasses) ou en **transition de phase** (phénomènes de bifurcation, émergence de clusters). Le **DSL**, en se concevant comme un **système** de variables $\omega_{i,j}$ s’ajustant localement, fournit un terrain d’étude où l’on peut appliquer des outils de **physique statistique** (recuit simulé, énergie libre, etc.) pour analyser la formation de **macro-clusters** et les “sauts” de configuration. On peut, en outre, importer dans le DSL des idées de frustration ou d’ordre–désordre propre à ces modèles, afin de mieux comprendre les **changements de régime** dans le réseau.

En **sociologie** et en **économie**, on observe la constitution de **réseaux** humains, qu’ils concernent des relations d’amitié, des collaborations d’équipes ou des alliances politiques. Les principes du **DSL**, qui renforcent les **liaisons** lorsque la **synergie** est élevée et les affaiblissent dans le cas contraire, rappellent certaines **lois sociologiques**. L’**effet Matthieu**, souvent formulé par “les

riches s'enrichissent", illustre cette dynamique où un **nœud** bien connecté a plus de chances d'attirer de nouvelles connexions, renforçant ainsi sa centralité dans le réseau. Appliquer un cadre de **DSL** à des nœuds représentant des acteurs (individus, entreprises) peut éclairer la dynamique de création et de dissolution de liens, voire la **formation de coalitions** ou de groupes d'intérêt. À l'inverse, la sociologie propose des **règles** comportementales (incitations, répulsions, normes sociales) qui peuvent être injectées dans la définition de la **synergie** $S(\mathcal{E}_i, \mathcal{E}_j)$ ou dans l'**énergie** globale \mathcal{J} . Cela établit une **interaction** interdisciplinaire où la **sociologie** propose des principes de dynamique sociale, tandis que le **DSL** en modélise la **co-évolution** sous forme d'un réseau adaptatif de pondérations.

1.8.6.3. Avantages pour la Modélisation et la Simulation

Le **Deep Synergy Learning (DSL)**, outre son rôle au sein de l'IA, propose un **cadre** unificateur pour comprendre et modéliser l'**auto-organisation** dans de nombreux champs. Le **modèle d'actualisation** des pondérations $\{\omega_{i,j}\}$ suit la règle :

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta[S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)]$$

Cette dynamique possède une **portée interdisciplinaire**. Selon le contexte, la synergie $S(\mathcal{E}_i, \mathcal{E}_j)$ peut représenter un **gain écologique** entre espèces coexistant, un **bénéfice économique** entre partenaires commerciaux ou une **coopération sociale** entre agents interagissant dans un réseau adaptatif. Il devient alors possible de **simuler** un vaste éventail de comportements collectifs, de détecter l'émergence de **communautés** (clusters) et d'analyser la **stabilité** ou les **transitions** (par exemple, la soudaineté avec laquelle le réseau se réorganise en une configuration inédite). Cette vision évoque le travail en **physique statistique** ou en **théorie des réseaux complexes**, où l'on s'intéresse à la dynamique d'un grand nombre d'éléments en interaction.

Un **avantage** majeur dans cette optique est la **dimension expérimentale** du DSL. Comme il se prête naturellement à une **implémentation** computationnelle, chaque hypothèse sur la fonction $S(\mathcal{E}_i, \mathcal{E}_j)$, le paramétrage η, τ , ou encore la structure initiale du réseau, peut être **testée** par simulation. Les résultats, qu'il s'agisse de la **constitution de clusters**, d'**oscillations**, de **convergence** ou de **fragmentation**, se **comparent** aux observations empiriques d'autres disciplines. En **écologie**, ils peuvent refléter la **dynamique des populations**, en **économie**, la **formation de coalitions**, et en **sociologie**, l'**essor ou le déclin de communautés**. Le **DSL** fonctionne alors comme un **laboratoire** virtuel où chaque domaine fournit ses **informations** ou **règles** pour définir la **synergie**. Le **SCN (Synergistic Connection Network)** en simule les **évolutions**, permettant d'observer comment les **entités** interagissent et se réorganisent au fil du temps.

Dans ce **dialogue**, il est possible de **valider** ou **invalidier** différentes hypothèses de fonctionnement collectif en examinant la façon dont les **pondérations** évoluent. Les disciplines comme la **physique des systèmes complexes** ou la **théorie des réseaux** y trouvent un nouvel objet de recherche. Le **DSL** propose un **réseau** dont les liens s'adaptent localement aux bénéfices mutuels des nœuds, au lieu de rester statiques ou de suivre une loi exogène. Cette boucle d'expérimentation renforce la **pertinence** du DSL en tant qu'instrument scientifique, au-delà de son utilité en ingénierie de l'intelligence artificielle.

1.8.6.4. Conclusion

Le **Deep Synergy Learning** ne se limite pas à l'IA ou au machine learning. Il s'intègre dans une **dynamique plus large** de **systèmes auto-organisés**, où les interactions locales façonnent l'évolution globale du réseau, rapprochant :

- Les **neurosciences** (assemblées neuronales, plasticité),
- Les **sciences cognitives** (cognition distribuée),
- L'**écologie** et la **physique des systèmes complexes** (réseaux dynamiques, transitions de phase),
- La **sociologie** et l'**économie** (formation de coalitions, interactions auto-renforcées).

Cette **interdisciplinarité** offre des **perspectives** nouvelles. Le **DSL** peut fournir un **cadre unifié** pour l'**auto-organisation** dans divers domaines, tout en intégrant des **principes** issus de ces disciplines, comme la **plasticité**, la **dynamique des systèmes**, la **parcimonie** ou encore les **mécanismes de contrôle**, dans l'évolution des pondérations ω . En somme, l'**effet** du DSL sur la recherche interdisciplinaire tient à la fois d'une **inspiration réciproque** (d'emprunts aux sciences complexes, biologiques, sociales) et d'une **offre** d'outils conceptuels pour simuler et analyser des **réseaux évolutifs**. C'est cette fertilisation croisée qui peut accélérer l'avancée vers des **systèmes** plus "vivants", plus "cognitifs" et plus adaptés à la complexité du monde réel.

1.8.7. Exemple de Convergence : DSL & Neurosciences

Dans les différentes sections de ce **chapitre 1.8**, qui explorent le **positionnement** du **Deep Synergy Learning (DSL)** par rapport aux paradigmes existants en IA et aux domaines connexes (IA symbolique, apprentissage profond, apprentissage par renforcement, logiques distribuées), un **fil conducteur** se dégage. Le **DSL** s'inspire fortement de la **biologie** et ses mécanismes d'**auto-organisation** présentent une **parenté** avec les **dynamiques neuronales** et d'autres phénomènes biologiques adaptatifs. De fait, un champ particulièrement fécond pour valider ou stimuler le DSL est celui des **neurosciences**. Dans cette sous-section (1.8.7), nous examinerons comment le **DSL** et la **neurobiologie** (ou les neurosciences cognitives) peuvent **converger** en partageant des hypothèses, des modèles et des interprétations. Une analogie peut être établie entre la **plasticité synaptique** (règle de Hebb) et la mise à jour des pondérations $\omega_{i,j}$ dans le **DSL**, où les connexions se renforcent ou s'affaiblissent selon leur synergie perçue. L'architecture du **DSL** pourrait également fournir un cadre pertinent pour modéliser l'**émergence** d'assemblées neuronales et la dynamique cérébrale, en capturant la formation spontanée de clusters coopératifs.

Les **neurosciences** offrent en retour des **principes** pouvant enrichir le **DSL**, notamment à travers des mécanismes comme la **saturation synaptique**, l'**inhibition compétitive** ou encore l'**organisation en couches corticales**. L'intégration de ces éléments pourrait rendre le DSL plus **biologiquement plausible**, tout en améliorant notre compréhension des processus cognitifs à travers un cadre mathématique et computationnel inspiré du fonctionnement du cerveau.

1.8.7.1. Règle de Hebb et Mise à Jour des Pondérations

Les **neurosciences** ont mis en évidence la **plasticité synaptique**, illustrée par la **règle de Hebb**. Lorsque des neurones sont activés simultanément de manière répétée, leurs **connexions** se renforcent. De manière analogue, dans le **DSL**, les pondérations $\omega_{i,j}$ évoluent en fonction de la **synergie** perçue entre entités. Cette loi fondamentale explique la formation d'**assemblées** neuronales et l'émergence de **mémoires** dans le cerveau. Le **Deep Synergy Learning (DSL)**, en mettant à jour les **pondérations** $\omega_{i,j}$ selon la dynamique :

$$\omega_{i,j}(t+1) = \omega_{i,j}(t) + \eta[S(\mathcal{E}_i, \mathcal{E}_j) - \tau \omega_{i,j}(t)],$$

prolonge directement ce principe. En effet, la fonction $S(\mathcal{E}_i, \mathcal{E}_j)$ tient lieu de mesure de **coopération** ou de **co-activation** (corrélation, co-information, etc.), et agit comme un **signal** incitant au **renforcement** local des liens. Lorsque deux entités \mathcal{E}_i et \mathcal{E}_j se trouvent *fréquemment* en situation d'apport mutuel, la pondération $\omega_{i,j}$ augmente, à la manière de deux neurones biologiques se “synchronisant” de plus en plus. Cette analogie étend l'**esprit** de la règle de Hebb à un cadre plus général. La **synergie** ne se limite plus à la simple co-activation instantanée, mais intègre des interactions **n-aires** et des mesures **entropiques** plus fines, reflétant des dynamiques complexes au sein du **DSL**.

Le terme $-\tau \omega_{i,j}(t)$ introduit un facteur de **décroissance** évitant que tous les liens ne croissent indéfiniment. Cela s'apparente à la **régulation** homéostatique neuronale, où le système limite les synapses trop puissantes pour maintenir un équilibre. De cette façon, le **DSL** hérite des principes de la **plasticité synaptique** et de la **règle de Hebb**, tout en les élargissant à un contexte d'**auto-organisation** plus souple (synergie multi-entités, possibilité de créer ou rompre massivement des liaisons, etc.). Ainsi, chaque pondération $\omega_{i,j}$ reflète une **dynamique** coopérative hébbienne, tout en gérant la **compétition** (diminution) nécessaire à la stabilité globale du réseau.

1.8.7.2. Assemblées Neuronales et Clusters dans le DSL

Un concept clé en **neurosciences cognitives** est celui d'**assemblée neuronale**, un ensemble de neurones interconnectés qui forme une configuration **stable** et se co-active lorsqu'un **concept**, un **souvenir** ou une **action** est représenté. Ces assemblées se créent, se renforcent et finissent par disparaître sous l'action de la **plasticité** synaptique, rendant compte de la manière dont le cerveau encode et renouvelle ses représentations.

Dans le **Deep Synergy Learning (DSL)**, l'on décrit de manière similaire la formation de **clusters** lorsque plusieurs entités $\{\mathcal{E}_k\}$ entretiennent entre elles des liaisons $\omega_{k,m}$ suffisamment élevées. L'existence de fortes **pondérations** indique que ces entités agissent conjointement et forment un **macro-cluster** relativement **stable**, lequel perdure tant que la **synergie** interne demeure supérieure aux forces de dispersion. Cette situation évoque la co-activation neuronale d'une assemblée, c'est-à-dire un **groupe** d'entités rassemblées pour traiter ou représenter une information.

La **dynamique** de ces clusters dans le DSL rappelle alors la **dynamique** des assemblées neuronales. Sous l'effet d'événements exogènes ou de signaux inattendus, certains liens $\omega_{k,m}$ peuvent se voir affaiblis ; le cluster se scinde et de nouveaux **sous-groupes** émergent, reflétant un changement de « contexte » ou de « contenu » que le réseau doit coder. Dans ce cadre, la flexibilité de l'**auto-organisation** (section 1.4.5) reflète la plasticité synaptique qui permet au cerveau de réorganiser ses **assemblées**. Un concept peut s'affaiblir et disparaître, tandis qu'un autre émerge en réponse à l'évolution des données perceptives. Cette **correspondance** entre les **clusters** du DSL et les **assemblées neuronales** biologiques suggère que le DSL pourrait apporter un **modèle** formel pour étudier la **cognition** ou la **perception** en flux continu, tout comme la plasticité cérébrale propose une source d'inspiration pour la gestion auto-adaptative des liens ω .

1.8.7.3. Mécanismes Biologiquement Plausibles et Leur Transposition

De nombreux modèles de **plasticité** neuronale insistent sur le fait que le cerveau ne se limite pas aux **synapses** excitatrices, qui consolident les connexions. Il recourt également à des **synapses** inhibitrices, introduisant un **frein** ou une **régulation** nécessaire pour éviter l'excitation excessive. Dans le cadre d'un **Deep Synergy Learning (DSL)**, il est possible de transposer cette idée en différenciant des liens de **coopération** et des liens d'**inhibition**. Les pondérations élevées traduisent une synergie positive entre deux entités, tandis que l'effet inhibiteur peut jouer un rôle similaire à la saturation ou à la normalisation compétitive, où l'entité est contrainte de répartir ses liaisons sur un volume total maximum. Cette approche s'accorde davantage avec la **régulation** rencontrée en biologie et empêche la dynamique de ω de s'emballer.

Dans un second temps, on peut s'inspirer de la **hiérarchie** observée dans le cortex biologique. Le cerveau se divise en aires sensorielles primaires et en aires associatives plus élaborées, reliées par des voies de feedforward, de feedback et latérales. Un **DSL** “multi-niveau” peut mimer cette structure, en répertoriant différents sous-ensembles d'entités en “couches” ou “aires” spécifiques. Chaque sous-ensemble d'entités s'auto-organise localement, tandis que des connexions inter-aires agissent comme des ponts entre ces zones. L'**auto-organisation** continue alors à se dérouler, mais la topologie n'est plus entièrement libre. Une **architecture** partielle est imposée pour refléter la **topographie** ou la **fonction** de certaines régions corticales. Cette hybridation, combinant une **contrainte** inspirée de la hiérarchie cérébrale et la **liberté** d'auto-organisation des liaisons ω , rend le DSL plus proche encore d'un **modèle** neuronal plausible. On peut ainsi concevoir un système qui, à la fois, bénéficie de la capacité plastique du DSL, tout en maintenant une division en zones (aires sensorielles, associatives, motrices) comme dans le cortex, assurant la **cohésion** et la **spécialisation**.

1.8.7.4. Apports Potentiels pour Comprendre le Cerveau

Le rapprochement entre les principes d'**auto-organisation** du **Deep Synergy Learning (DSL)** et ceux de la **plasticité neuronale** ne se limite pas à une simple analogie. Il ouvre des **perspectives** de recherche où un **réseau** DSL peut être utilisé pour **simuler** ou **reproduire** certains phénomènes neuronaux. Cette démarche peut s'articuler en plusieurs axes. Le **premier** consiste à implémenter un système DSL dans lequel on introduit des **stimuli** ou des signaux analogues à ceux que reçoit un cortex sensoriel, en observant la formation de **clusters** ou d'**assemblées** stables qui coderaient

ces stimuli. Il devient ainsi envisageable d'étudier, dans un cadre simulé, comment des ensembles de pondérations $\omega_{i,j}$ peuvent évoquer une **attention** sélective, un découpage temporel de l'information ou même une rémanence mnésique.

Un **second** volet réside dans la **comparaison** directe entre les **données** issues des simulations DSL et des **enregistrements** neuronaux réels (par exemple des signaux obtenus via EEG, MEG, fMRI, ou encore des relevés d'activité calcique à l'échelle microscopique). Il s'agirait d'examiner si des **patterns** d'activation ou de "clusters dynamiques" dans le réseau DSL reproduisent qualitativement ou quantitativement des signatures observées dans le cerveau. Les similitudes, si elles se révèlent robustes, conforteraient l'idée que la dynamique auto-organisée $\omega(t)$ peut effectivement mimer la constitution et la dissolution d'**assemblées** neuronales chez un organisme vivant.

Le **troisième** axe consiste à proposer des **hypothèses** sur la **plasticité** cérébrale en s'appuyant sur la structure du DSL. Il serait possible, par exemple, d'y implanter des **règles** inspirées de l'anti-Hebb, des mécanismes de **synaptic pruning** (réduction des synapses peu sollicitées), ou des saturations synaptiques plus sophistiquées, afin de **tester** leur impact sur la dynamique du réseau. Les résultats issus de ces **expérimentations informatiques** pourraient orienter des recherches en **neurosciences**. Si un mécanisme de **régulation** des pondérations ω améliore significativement la **stabilité** ou le **codage** dans le réseau DSL, on peut alors émettre l'**hypothèse** qu'un processus similaire pourrait exister dans le cerveau.

Au final, ce **dialogue** entre la simulation DSL et les **observations** neuronales enrichit à la fois la compréhension de la **plasticité** et des fonctions cognitives, tout en dotant le DSL de **mécanismes** biologiquement inspirés susceptibles d'améliorer son adaptabilité et sa capacité d'**apprentissage**. Les **prochains** développements, où l'on confronte plus systématiquement les modèles DSL aux données empiriques de neurosciences, pourraient rapprocher encore davantage cette classe de réseaux des **dynamiques** neuronales réelles, et peut-être contribuer à expliquer certains aspects de la mémoire, de la consolidation et de l'attention.

Conclusion

Le **Deep Synergy Learning** et les **neurosciences** peuvent converger de plusieurs manières :

- **Analogie** fondamentale entre la plasticité synaptique (Hebb) et la mise à jour $\omega_{i,j}$,
- **Formation** de clusters stables rappelant les "assemblées neuronales" du cerveau,
- **Inspiration** mutuelle pour implémenter des mécanismes biologiquement plausibles (inhibition, saturation, hiérarchies corticales),
- **Modélisation** de phénomènes cognitifs ou neuronaux complexes, offrant aux neurosciences un outil de simulation distribué, et recevant en retour des insights pour stabiliser et faire évoluer le DSL.

Cette **relation** avec les neurosciences illustre l'idée d'un paradigme **interdisciplinaire** (voir 1.8.6) où le DSL agit comme un **pont** entre l'IA et la biologie, susceptible d'aider à **comprendre** la cognition naturelle tout en **étendant** les capacités adaptatives des systèmes artificiels.

