编译神经网络为加权自动机：一个统一的形式化分析框架

1. 理论基础：通用的溯源代数框架

本部分介绍作为整个研究基石的通用代数工具与核心思想，它们是后续具体应用于不同网络结构的基础

1.1. 问题的提出与代数障碍

1️⃣ **分段线性(PWL)特性**: 一个包含ReLU激活函数的深度神经网络，其本质是一个高维空间中的**连续分段线性函数(Continuous Piecewise Linear, PWL)** $F : \mathbb{R}^n \to \mathbb{R}^m$

2️⃣ **代数障碍**: ReLU函数 $\mathrm{ReLU}(z) = \max(0, z)$ 的非线性特性，构成了形式化分析的主要障碍 具体而言：

1. 它使得传统的**线性空间(Linear Space)**理论无法直接、完整地描述整个网络
2. 它将函数从简单的线性或多项式形式转化为更复杂的**分段线性(PWL)**形式，导致**多项式环(Polynomial Ring)**等经典代数结构在逐层计算中失效

1.2. 核心方法论：交换半环与数据溯源

🆗 **交换半环 (Commutative Semiring)**: 介绍交换半环 $(K, \oplus, \otimes, \mathbf{0}_K, \mathbf{1}_K)$ 的概念 与环(Ring)相比，半环不要求加法逆元（即减法），使其能更灵活地对如 `max` 等非标准运算进行建模 这对于捕捉ReLU激活的本质至关重要

1.3. 核心代数框架的构建

1️⃣ **指导思想**: 将网络中"线性"与"非线性"的计算，解耦并映射到两个不同的、代数上完备的结构中，再通过一个投影算子建立联系

2️⃣ **核心结构定义**:

1. **预激活空间 (Space of Pre-activations)** $\mathcal{P}$: 一个由输入变量 $\mathbf{x} \in \mathbb{R}^n$ 上的**连续分段线性函数 (PWL) 构成的向量空间 (Vector Space)**
   - **作用**: 专门处理网络中的所有线性运算（如加权求和） 由于PWL函数的线性组合（函数加法与标量乘法）的结果仍然是PWL函数，该空间在此类运算下是**封闭的**
2. **激活半环 (Semiring of Activations)** S_+: 一个由**非负**连续分段线性函数构成的交换半环
   - **作用**: 专门容纳经过ReLU等非负激活函数之后的结果 其代数结构（如将$\max$作为加法）忠实地反映了激活函数的计算特性 重要的是，S_+ 中的每个元素本身也属于 $\mathcal{P}$ 空间
3. **投影算子 (Projection Operator)** σ: 一个从预激活空间到激活半环的映射，即 $\sigma : \mathcal{P} \to S_+$
   - 最典型的即 **ReLU投影算子**: $\sigma_{\mathrm{relu}}(P_z) := \max(0, P_z)$，它将一个任意的PWL函数映射为一个非负的PWL函数

1.4. 该代数体系的"适用性边界"清单

组件	完全适用	不适用/需近似
激活函数	ReLU，Leaky ReLU，PReLU，Identity，Absolute Value	Sigmoid，Tanh，Softmax，ELU，GELU， Swish，etc.

组件	完全适用	不适用/需近似
线性层	全连接（Dense），卷积（Convolutional）	—
池化/翼合	最大池化（Max Pooling），平均池化（Average Pooling），和/均值/最大聚合	—
归一化层	推理模式下的批量归一化（Inference—mode BatchNorm）	训练模式的BN，LayerNorm，GroupNorm， etc.
交互门控	加法/残差连接（Additive Skip—connections）	乘法门控（LSTM/GRU gates），自注意力 （Self—Attention）
结构/拓扑	有向无环图（DAGs），FFN，CNN，ResNet，固定图上的GNN	循环结构（RNNs），动态图（Dynamic Graphs）
其他	—	Dropout（随机性）

2. 从神经网络到溯源代数

2.0. 用溯源代数描述三类典型的神经网络

1️⃣ FFN的溯源推导算法

1. **初始化**: 输入层 $l=0$ 的激活溯源定义为输入变量本身，$P(a_i^{(0)}) = x_i$ 由于线性函数是最简单的PWL函数，此表达式属于预激活**空间** $\mathcal{P}$
2. **逐层传播**: 对于第 l ($l \geq 1$) 层的神经元 j:
   - (a) 在 $\mathcal{P}$ 中计算预激活溯源: $P(z_j^{(l)}) = \sum_i w_{ji}^{(l)} \cdot P(a_i^{(l-1)}) + b_j^{(l)}$
   - (b) 投影到 S_+ 中计算激活溯源: $P(a_j^{(l)}) = \sigma(P(z_j^{(l)}))$
3. **最终输出**: 网络最后一层的 $P(a_j^{(L)})$ 或其线性组合 $P(\mathbf{y})$

2️⃣ 针对CNN的代数框架扩展

0. 为了将溯源代数应用于CNN，我们需将CNN的核心算子映射到我们的代数结构 $(\mathcal{P}, S_+, \sigma)$ 中
1. **符号的张量化**: 首先，将溯源表达式从向量索引 $P(a_i^{(l)})$ 扩展为携带空间和通道信息的张量索引 $P(a)_{l,k,i,j}$，以适应图像数据的结构
2. **线性算子的映射 (卷积与平均池化)**:
   - **卷积 (Convolution)**: 本质上是一系列带权重共享的局部线性组合
   - **平均池化 (Average Pooling)**: 是一种无权重的局部线性组合
   - 由于我们的预激活空间 $\mathcal{P}$ 是一个对线性组合封闭的PWL函数向量空间，这两个关键的线性算子可以被完美地在 $\mathcal{P}$ 中建模，其结果仍为PWL函数
3. **非线性算子的映射 (最大池化)**:
   - **最大池化 (Max Pooling)**: 该操作在代数上与ReLU激活函数的核心运算——取最大值——是同构的 因此，它可以被直接建模为激活半环 S_+ 中的广义加法运算 $\bigoplus$ 一个N元最大池化等价于一个N元的半环加法

3️⃣ 针对GNN的代数框架扩展

1. **核心挑战与前提**: 与CNN中固定的邻域结构不同，GNN的计算依赖于输入的、可变的图拓扑结构 $G=(V,E)$ 这为建立通用的符号表达式带来了挑战 因此，本框架的核心应用前提是：**将图的拓扑结构 $G=(V,E)$ 视为固定的、给定的先验知识** 分析将在单个、具体的图实例上进行，以该图的邻接关系为基础，推导关于节点/图特征的符号表达式
2. **溯源符号的节点化**: 将溯源表达式与图中的节点和层数进行绑定 一个节点的特征向量的溯源被表示为 $P(h_v^{(l)})$，其中 $v \in V$ 是节点索引，l 是GNN的层数
3. **GNN核心运算的代数映射**:
   - **聚合(Aggregation)阶段**:
     - 对于**和聚合/均值聚合 (Sum/Mean Aggregation)**: 邻居信息的线性组合可以直接在预激活**空间** $\mathcal{P}$ 中表示 例如，对于和聚合，其溯源为 $\sum_{u\in\mathcal{N}(v)} P(h_u^{(l-1)})$，此运算在 $\mathcal{P}$ 向量空间中是封闭的
     - 对于**最大聚合 (Max Aggregation)**: 该操作与CNN的最大池化在代数上同构，可以直接应用泛化后的激活半环 S_+ 及其N元$\max$算子（即广义半环加法 $\bigoplus$）进行建模
   - **更新(Update)阶段**: GNN的更新步骤通常是应用一个共享的MLP（前馈网络） 这可以直接套用FFN的溯源推导算法

2.1. 代数框架的良定义性

✅ **定理FFN-1：代数框架的良定义性**

1. **陈述**: 对于任何ReLU网络，其任意激活节点的溯源表达式 $P(a_j^{(l)})$ $(l \geq 1)$ 都是激活半环 S_+ 中的一个合法的非负函数元素
2. **意义**: 保证了推导过程在代数上是**封闭且一致的** 线性组合的封闭性由 $\mathcal{P}$ 的向量空间性质保证，而非负激活的封闭性由投影算子 σ 和半环 S_+ 的定义保证

✅ **定理CNN-1：代数框架的良定义性**

1. **陈述**: 对于任何由卷积、池化、ReLU和全连接层构成的CNN，其任意节点的溯源推导算法都是**良定义的(well-defined)**，其结果在预激活空间 $\mathcal{P}$ 和激活半环 S_+ 中保持封闭
2. **意义**: 保证了CNN所有核心算子都能被统一的代数框架精确建模，是所有后续CNN形式化分析的**封闭性基石**

✅ **定理GNN-1： 代数框架的良定义性**

1. **陈述**: 对于一个给定的图 $G=(V,E)$ 和任何由消息传递层构成的GNN，其溯源推导算法是**良定义的(well-defined)**，其结果在预激活空间 $\mathcal{P}$ 和激活半环 S_+ 中保持封闭
2. **意义**: 确保了对GNN的符号分析在其**特定的图上**是代数自洽的，是所有后续GNN形式化分析的理论前提

2.2. 同态求值定理

✅ **定理FFN-2：同态求值定理**

1. **陈述**: 存在一个求值同态 $h_{\mathbf{x}_0}: S_+ \to \mathbb{R}$ 对于任意具体的输入向量 $\mathbf{x}_0$，以下等式成立：

$$h_{\mathbf{x}_0}(P(\mathbf{y})) = F(\mathbf{x}_0)$$

2. **意义**: 保证了我们的符号解释与真实网络的数值计算精确对应

✅ **定理CNN-2： CNN的同态求值定理**

1. **陈述**: 存在一个求值同态 $h_{\mathbf{x}_0}: S_+ \to \mathbb{R}$，使得对于任意输入张量 $\mathbf{x}_0$，以下等式恒成立：

$$h_{\mathbf{x}_0}(P(\mathbf{y})) = F(\mathbf{x}_0)$$

2. **意义**: 保证了符号分析是对真实网络计算行为的**忠实、精确**的数学建模，是连接符号世界与数值世界的**忠实性桥梁**

✅ **定理GNN-2： GNN的同态求值定理**

1. **陈述**: 存在一个求值同态 $h_{(\mathbf{X}_0)}$，使得对于任意输入节点特征矩阵 $\mathbf{X}_0$ 和图 G，以下等式恒成立：

$$h_{(\mathbf{X}_0)}(P(\mathbf{y})) = F(\mathbf{X}_0, G)$$

2. **意义**: 保证了GNN的符号分析能够**精确对应**其在特定图上的真实计算行为，为所有形式化分析提供了最终的现实依据

2.3. 结构性质定理

✅ **定理FFN-3：结构性质定理（连续性、分段线性与条件凸性）**

1. **陈述**:
   - (通用性质) 对于任意单输出ReLU网络，其溯源表达式 $P(y)$ 是一个定义在 $\mathbb{R}^n$ 上的**连续、分段线性**函数
   - (条件性质) 若该网络的所有**权重均为非负** ($w_{ji}^{(l)} \geq 0$)，则 $P(y)$ 同时也是一个**凸函数**
2. **意义**: 该定理的结论是溯源代数框架**内禀属性的直接体现** 它从代数构造的层面直接确认了ReLU网络函数的通用几何本质，并指出了其在特定条件下具备的更强属性（凸性），为优化和决策边界分析提供了理论基础

✅ **定理CNN-3：CNN的结构性质定理**

1. **陈述**:
   - (通用性质) 对于任何适用的单输出CNN，其输出溯源 $P(y)$ 是一个关于输入像素 $\mathbf{x}$ 的**连续、分段线性(PWL)** 函数
   - (条件性质) 若该CNN的所有权重均为非负，则 $P(y)$ 同时也是一个**凸函数**
2. **意义**: 确认了CNN所计算函数的基本数学性质，是所有基于梯度和几何形式化分析的理论基础

✅ **定理GNN-3：GNN的结构性质定理**

1. **陈述**:
   - (通用性质) 对于给定的图 G 和任何适用的单输出GNN，其输出溯源 $P(y)$ 是一个关于输入节点特征 $\mathbf{X}$ 的**连续、分段线性(PWL)** 函数
   - (条件性质) 若该GNN的所有权重均为非负，则 $P(y)$ 同时也是一个**凸函数**
2. **意义**: 将核心结构性质推广至图数据，确立了GNN函数在节点特征空间中的基本几何形态

2.4. 符号可微性与梯度定理

✅ **定理FFN-4：符号可微性与梯度定理**

1. **陈述**: 函数 $P(y)$ 在 $\mathbb{R}^n$ 上几乎处处可微，其梯度 $\nabla P(y)$ 是一个分段常数向量函数，并且可以通过对 $P(y)$ 进行符号运算直接求出
2. **意义**: 为实现精确的、非近似的梯度特征归因方法（如Saliency Map）提供了可能

✅ **定理CNN-4：CNN的符号可微性与梯度定理**

1. **陈述**: 作为 `定理CNN-3` 的推论，CNN的输出溯源函数 $P(\mathbf{y})$ 在其定义域上几乎处处可微，其梯度 $\nabla_{\mathbf{x}} P(\mathbf{y})$ 是一个分段常数张量函数，可通过符号运算直接求出
2. **意义**: 提供了计算CNN精确、无近似的特征归因（如Saliency Map）的引擎，可生成对一个输入区域都有效的归因函数

✅ **定理GNN-4：GNN的符号可微性与梯度定理**

1. **陈述**: 作为 `定理GNN-3` 的推论，GNN的输出溯源函数 $P(\mathbf{y})$ 是关于输入节点特征 $\mathbf{X}$ 几乎处处可微的，其梯度 $\nabla_{\mathbf{X}} P(\mathbf{y})$ 可通过符号求导获得
2. **意义**: 赋能GNN的高保真可解释性，允许精确计算每个输入节点特征对最终预测的贡献度

2.5. 活性分析定理

✅ **定理FFN-5：神经元活性分析定理**

1. **陈述**: 若能证明某神经元的预激活溯源 $P(z_j^{(l)})(\mathbf{x}) \leq 0$ 在整个输入域上成立，则该神经元是"死亡神经元"
2. **意义**: 提供了通过形式化方法精确识别冗余神经元的理论依据，可用于有保证的网络剪枝

✅ **定理CNN-5：滤波器活性分析定理**

1. **陈述**: 对于第 l 层的第 k 个滤波器，若能证明其预激活溯源 $P(z)_{l,k,i,j}$ 在**所有**空间位置 (i,j) 上，对于**整个**合法的输入域 $\mathcal{D}_{in}$，都满足 $P(z)_{l,k,i,j}(\mathbf{x}) \leq 0$，则该滤波器是一个"死亡滤波器"(Dead Filter)
2. **意义**: 将CNN的结构化剪枝问题从经验性评估，转化为一个可被形式化方法求解的**数学证明问题**，为实现有保证的网络压缩提供了理论依据

✅ **定理GNN-5：共享权重活性分析定理**

1. **陈述**: 对于GNN第 l 层更新函数（一个共享的MLP）中的第 j 个神经元，若能证明其预激活溯源 $P(z_j^{(l)})$，在**图G的所有节点** $v \in V$ **上**，对于**整个**合法的输入域 $\mathcal{D}_{in}$，都满足 $P(z_j^{(l)})(\mathbf{x}_v, \mathbf{m}_v) \leq 0$（其中$\mathbf{m}_v$为聚合信息），则该神经元对应的权重在**整个图**的计算中是"死亡"的
2. **意义**: 将GNN的模型压缩问题形式化 它为**结构化地剪枝GNN的共享权重**提供了数学保证，比基于采样的剪枝方法更为可靠和高效

2.6. 局部鲁棒性形式化验证定理

✅ **定理FFN-6：局部鲁棒性形式化验证定理**

1. **陈述**: 给定输入 $\mathbf{x}_0$ 和邻域 $B(\mathbf{x}_0, \epsilon)$，网络的输出在邻域内是否保持某性质，可以通过分析溯源表达式 $P(\mathbf{y})$ 在约束区域上的值域来精确判定 **
2. **意义**: 将鲁棒性验证问题从大量的抽样测试，转变为一个可以被自动化形式化工具求解的符号分析问题

✅ **定理CNN-6：CNN局部鲁棒性形式化验证定理**

1. **陈述**: 给定输入图像 $\mathbf{x}_0$ 和 ϵ-邻域 $B_p(\mathbf{x}_0, \epsilon)$，网络在该邻域内的鲁棒性，可通过分析其溯源表达式 $P(\mathbf{y})$ 在该约束区域上的值域来被精确判定
2. **意义**: 将CNN的对抗攻击鲁棒性验证问题，从经验性、基于采样的测试，转变为可被形式化工具精确求解的符号分析问题，为提供可证明的(provable)安全保证奠定基础

✅ **定理GNN-6：节点特征鲁棒性的形式化验证定理**

1. **陈述**: 给定图 G 和输入特征 $\mathbf{X}_0$，对于某个节点的预测，其在一个以 $\mathbf{X}_0$ 为中心、半径为 ϵ 的邻域 $B(\mathbf{X}_0, \epsilon)$ 内的稳定性，可以通过分析输出溯源表达式 $P(\mathbf{y})$ 在该约束区域上的值域来精确判定

2. **意义**: 将GNN**对节点特征的对抗攻击**的鲁棒性验证问题，从依赖大量对抗样本的经验性测试，转变为一个可以被自动化形式化工具求解的符号约束满足问题 这对于GNN在安全攸关领域的应用（如金融风控、社交网络机器人检测）至关重要

2.7. 组合性定理

✅ **定理FFN-8：分层溯源与线性组合性定理**

1. **陈述**: 任意一个神经元在第 l 层的激活溯源 $P(a_j^{(l)})$，可以被精确表示为前一层（$l-1$ 层）所有激活溯源 $\{P(a_i^{(l-1)})\}$ 的一个连续、分段线性的符号函数
2. **意义**: 该定理在数学上精确刻画了FFN的层级计算结构 它将整个网络的复杂函数分解为一系列更简单的、逐层复合的PWL函数，是进行层级化分析与理解FFN特征抽象过程的理论基石

✅ **定理CNN-7：分层溯源与特征组合性定理**

1. **陈述**: 任意一个激活节点的溯源表达式 $P(a)_{l,k,i,j}$，可以表示为前一层激活 $P(a)_{l-1,:,:,:}$ 的一个连续、分段线性的符号函数
2. **意义**: 在数学上**精确刻画了**CNN的层级计算机制，使得层与层之间的归因分析成为可能，为揭示CNN的**层级组合性**本质提供了终极的白盒工具

✅ **定理GNN-6：基于图的分层溯源与消息组合性定理**

1. **陈述**: 任意一个节点 v 在 l 层的激活溯源 $P(h_v^{(l)})$，可以表示为其在 $l-1$ 层的自身激活溯源 $P(h_v^{(l-1)})$ 以及其所有邻居节点激活溯源 $\{P(h_u^{(l-1)})|u \in \mathcal{N}(v)\}$ 的一个连续、分段线性的符号函数
2. **意义**: 在数学上**忠实地刻画了消息传递机制**的本质 它提供了一个形式化的白盒工具，使我们能够精确地、符号化地回溯一个节点的最终表示是如何由其多跳邻居的初始特征逐层、非线性地组合而来的，为理解GNN的"感受野"和信息流动提供了坚实的理论基础

2.8. 性形式化验证定理

✅ **定理FFN-8：输入特征对称性形式化验证定理**

1. **陈述**: 令 T 为一个作用于输入空间 $\mathbb{R}^n$ 的特定变换 网络 F 具备关于变换 T 的不变性，当且仅当以下符号恒等式能够被代数方法证明：

$$P(\mathbf{y}|T(\mathbf{x})) \equiv P(\mathbf{y}|\mathbf{x})$$

2. **意义**: 为验证FFN是否学习到了任何预期的（或非预期的）输入特征对称性提供了形式化工具，是进行模型调试、发现意外偏见和验证特定设计假说的重要手段

✅ **定理CNN-8：几何等变性的形式化验证定理 (Formal Verification of Geometric Equivariance)**

1. **陈述**: 令 $\mathcal{T}$ 为一个作用于输入空间的几何变换群（如平移），$\mathcal{T}'$ 为其在输出空间对应的变换群 网络 F 具备 $(\mathcal{T}, \mathcal{T}')$ 等变性，当且仅当对于代表变换的符号算子 $T \in \mathcal{T}$ 和 $T' \in \mathcal{T}'$，符号恒等式

$$P(\mathbf{y}|T(\mathbf{x})) \equiv T'(P(\mathbf{y}|\mathbf{x}))$$

   能够被代数方法证明

2. **意义**: 将网络的几何性质分析从不完备的经验性测试，转变为一个**单次的、完备的代数证明问题**，极大地提升了对模型内在性质分析的可靠性与效率

✅ **定理GNN-8：置换等变性的形式化验证定理 (Formal Verification of Permutation Equivariance)**

- **陈述**: 令 π 为一个作用于图节点顺序的置换算子，π' 为其在输出空间对应的置换算子 一个GNN模型 F 具备置换等变性，当且仅当对于代表置换的符号算子 π 和 π'，符号恒等式

$$P(\mathbf{y}|\pi(X,A)) \equiv \pi'(P(\mathbf{y}|X,A))$$

  能够被代数方法证明

- **意义**: 将GNN最核心的内在几何性质——**置换等变性**，从一个需要通过实验观察和归纳的经验属性，提升为一个可以通过**单次、完备的代数证明**来一劳永逸验证的数学定理 这极大地增强了对GNN模型可靠性分析的深度和广度

2.9. 符号化因果干预与反事实分析定理

✅ **定理FNN-9：符号化因果干预与反事实分析定理**

1. **陈述**: 对于一个由ReLU FFN F 导出的溯源表达式 $P(\mathbf{y})$，本框架能确定性地构造出在**干预算子** $do(a_k^{(l)} \leftarrow P_c)$ 作用下的**反事实溯源表达式**

$$P' \triangleq P(\mathbf{y}|do(a_k^{(l)} \leftarrow P_c))$$

   该表达式 P' 及其与原表达式的差（即**个体因果效应** $P_{\text{ICE}} \triangleq P' - P$），均为代数上良定义的连续分段线性(PWL)函数

2. **意义**: 该定理将分析能力从观察性的“关联分析”（如梯度）提升至干预性的“**因果分析**” 它通过构建形式化的反事实“what-if”场景，能够精确揭示网络内部组件的真实贡献，为实现更鲁棒的可解释性(XAI)、高精度模型调试与公平性审计提供了核心理论工具

✅ **定理CNN-9：结构化因果干预与特征图反事实分析定理**

1. **陈述**: 对于一个CNN模型 F 及其输出溯源表达式 $P(\mathbf{y})$，本框架能确定性地构造在**结构化干预算子** $do(P(a)_{l,k,:,:} \leftarrow P_{C,k})$ 作用下的**反事实溯源表达式**:

$$P' \triangleq P(\mathbf{y} \mid \text{do}(P(a)_{l,k,:,:} \leftarrow P_{C,k}))$$

   该表达式 P' 及其与原表达式的差（即**结构化个体因果效应** $P_{\text{S-ICE}} \triangleq P' - P$），均为代数上良定义的连续分段线性(PWL)函数

2. **意义**: 该定理将因果分析的粒度从单个神经元提升至语义上更关键的**整个特征图（滤波器）** 它允许对网络中的**抽象视觉概念**（如“纹理”或“眼睛”）进行直接干预，为实现形式化的**滤波器消融研究、基于概念的调试**与**高可信度视觉解释（XAI）** 提供了严谨的数学工具

✅ **定理GNN-9：共享特征维度的因果干预与反事实分析定理**

1. **陈述**: 设 F 为在图 $G=(V,E)$ 上定义的GNN模型 对于第 l 层共享的第 k 个特征维度，本框架能确定性地构造在以下**干预算子**作用下的**反事实溯源表达式** P':

$$P' \triangleq P\left(\mathbf{y} \mid \text{do}\left(P(h_{v,k}^{(l)}) \leftarrow P_{C,k} \quad \forall v \in V\right)\right)$$

   该表达式 P' 及**节点特征因果效应** $P_{\text{NF-ICE}} \triangleq P' - P$，均为代数上良定义的**连续分段线性(PWL)函数**

2. **意义**: 该定理提供了一种**形式化**的“**概念手术刀**”，使我们能精确剔除或改变GNN在整个图上学到的某个共享抽象概念（如节点的“中心性”角色），并量化这一“手术”对模型最终决策的真实影响

2.10. 形式化修复问题构建定理

✅ **定理FFN-10：FFN的形式化修复问题构建定理**

1. **陈述**: 设 F_W 为一个由权重 W 参数化的FFN模型，Φ 为一个该网络当前不满足的形式化性质 寻找一个最小的权重扰动 ΔW 以修复该网络的问题，可以被确定性地构建为以下关于变量 ΔW 的优化问题：

$$\begin{aligned} \underset{\Delta W}{\text{minimize}} \quad & \|\Delta W\|_p \\ \text{subject to} \quad & \text{Func}(P_{W+\Delta W}) \text{ satisfies } \Phi \end{aligned}$$

   其中，函数 $\text{Func}(P_{W+\Delta W})$ 是由溯源表达式 $P_{W+\Delta W}$ 所代表的数学函数 关键在于，约束条件 `satisfies` Φ 可被转化为一个关于 ΔW 中元素的**半代数集 (semi-algebraic set)**，即一组关于 ΔW 的多项式等式与不等式

2. **意义**: 该定理将FFN的"模型修复"问题，从启发式探索，转化为可在**实代数几何(Real Algebraic Geometry)**框架下进行分析的精确优化问题 它为开发基于符号计算或数值优化的、有保证的模型修复算法铺平了道路

✅ **定理CNN-10：CNN的形式化修复问题构建定理**

1. **陈述**: 设 F_W 为一个由权重 W 参数化的CNN模型，Φ 为一个当前被违反的形式化性质（例如，由`定理 CNN-8`所定义的几何等变性） 寻找最小权重扰动 ΔW 以修复该CNN的问题，可以同样被构建为一个约束优化问题 其约束可被转化为一个关于变量 ΔW 的半代数集

2. **意义**: 为形式化地修复CNN的深层属性（如对抗鲁棒性、几何不变性）提供了严谨的数学框架 它使得我们可以超越数据驱动的防御，从模型权重的层面直接进行**有数学保证的"外科手术式"修复**

✅ **定理GNN-10：GNN的形式化修复问题构建定理**

1. **陈述**: 设 F_W 为一个由权重 W 参数化的GNN模型，Φ 为一个当前被违反的形式化性质（例如，由`定理 GNN-8`所定义的置换等变性） 寻找最小权重扰动 ΔW 以修复该GNN的问题，可以同样被构建为一个约束优化问题 其约束可被转化为一个关于变量 ΔW 的半代V数集

2. **意义**: 将GNN的修复能力从节点或边的层面，提升至模型**共享权重**的层面 它为解决由训练偏差导致的系统性问题（如对特定社群的偏见，违反图的对称性）提供了根本性的、可计算的解决方案

3. 基于加权自动机的神经网络编译

3.1. 动机：从符号分析到自动机编译

1️⃣ **符号爆炸的必然性**: 第二章所构建的溯源表达式 $P(\mathbf{y})$，虽然在数学上是精确的，但其嵌套的树状结构对于深度和宽度较大的网络会呈指数级增长 这导致直接存储、操作和求解该符号表达式的开销变得无法承受，严重制约了理论框架的实用性

2️⃣ **解决方案的愿景**: 引入加权自动机——一种基于图的计算模型 其图结构天然支持对共享计算路径的合并，能够以远比符号树更紧凑、高效的方式来表示和操作复杂的PWL函数

3️⃣ **本章目标**: 构建一个以溯源代数为理论核心的**通用神经网络编译器** C 其目标不再是"证明"符号表达式和一个自动机等价，而是实现一个直接的**表示法转换**：将第二章生成的、树状的PWL函数符号表达式 $P(\mathbf{y})$，高效地编译为计算上等价、结构上紧凑的加权自动机 A_F：

$$P(\mathbf{y}) \xrightarrow{\text{Compiler } C} A_F, \quad \text{使得 } A_F(\mathbf{x}) \equiv h_{\mathbf{x}}(P(\mathbf{y})) \equiv F(\mathbf{x}) \quad \forall \mathbf{x} \in \mathcal{D}_{in}$$

3.2. 理论基础：PWL函数与加权自动机的内在等价性

3.2.1. 结构化输入的加权自动机

1️⃣ **形式化定义 (泛化)**: 一个加权自动机 A 是一个元组 $A = (Q, \Sigma, \delta, I, T, K)$，其中：

1. Q 是有限的状态集合
2. Σ 是输入**结构单元**的字母表 对于FFN，Σ 是特征向量；对于CNN，可以是像素网格；对于GNN，可以是节点和边
3. δ 是带权重的转移函数 其定义依赖于输入结构，例如 $\delta : Q \times \Sigma_{patch} \to K \times Q$ 用于CNN，或 $\delta : Q^{|\mathcal{N}(v)|} \times \Sigma_{node} \to K \times Q$ 用于GNN
4. I, T 分别是初始与终止权重函数
5. $K = (S, \oplus, \otimes, \mathbf{0}, \mathbf{1})$ 是一个**半环 (Semiring)**

3.2.2. 代数同源性：(max,+)代数与热带半环

1️⃣ **核心洞察**: 我们在第二章构建的整个溯源框架，其核心运算可以归结为**线性组合**（属于$\mathcal{P}$空间）和**取最大值**（属于S_+半环） 一个ReLU网络的最终函数本质上是一个由多个线性函数通过`max`和`+`组合而成的复杂函数 这种 $(\max, +)$ 计算结构，与著名的**热带半环 (Tropical Semiring)** $(\mathbb{R} \cup \{-\infty\}, \max, +)$ 在代数上是完全同构的 加权自动机正是运行在这种半环上的理想计算模型

3.2.3. 核心等价性定理

✅ **定理WFA-1：通用溯源-自动机等价性定理 (General Provenance-Automaton Equivalence Theorem)**

1. **陈述**: 对于任何由FFN、CNN或GNN（基于给定图拓扑）的溯源代数框架所生成的输出溯源表达式 $P(\mathbf{y})$，该PWL函数可以被一个与之计算等价的、可能接受结构化输入的加权自动机 A_P 所**表示** 该自动机的权重半环 K 与溯源代数的激活半环 S_+ (即 $(\max, +)$ 代数) 在计算上是兼容的

$$h_{\mathbf{x}_0}(P(\mathbf{y})) = A_P(\mathbf{x}_0)$$

2. **意义**: 它为神经网络的"编译"提供了理论合法性，确保了将复杂、低效的符号表达式转化为高效的加权自动机是数学上完全等价、无损的

3.3. 通用编译算法：神经网络的逐层自动机构建

1️⃣ 编译原则：算子的复合

1. 令 A_{l-1} 为代表网络前 $l-1$ 层所计算的PWL函数的自动机，编译第 l 层的过程，就是构造一个代表该层运算的变换算子 $\mathcal{T}_l$，并将其应用于 A_{l-1}，得到新的自动机 $A_l = \mathcal{T}_l(A_{l-1})$

2️⃣ 编译基础线性层 (FFN)

1. **操作**: $W\mathbf{a} + \mathbf{b}$，这是在 $\mathcal{P}$ 空间中的线性变换
2. **自动机变换** $\mathcal{T}_{linear}$: 此变换通过一个**加权变换器 (Weighted Transducer)** 与前层自动机 A_{l-1} 进行复合来实现 该变换器精确地将 A_{l-1} 所表示的PWL函数 $P(\mathbf{a})$ 映射为新的PWL函数 $W \cdot P(\mathbf{a}) + \mathbf{b}$

3️⃣ 编译ReLU激活层

1. **操作**: $\mathrm{ReLU}(z) = \max(0, z)$，这是从 $\mathcal{P}$ 到 S_+ 的投影 σ

2. **自动机变换 $\mathcal{T}_{relu}$**: 此变换通过对自动机的**状态分裂 (State Splitting)** 或其他等价操作来实现 它将代表 $P(z)$ 的自动机，转换为代表 $\max(0, P(z))$ 的新自动机，这在代数上对应于应用热带半环的加法（max）运算

4️⃣ 针对CNN的编译扩展

1. **卷积层**: 其**滑窗**和**权重共享**的线性运算，可被建模为一个结构化的**网格变换器 (Grid Transducer)**，它在本质上是 $\mathcal{T}_{linear}$ 的一种高效实现
2. **池化层**: **平均池化**作为线性操作，采用 $\mathcal{T}_{linear}$ 的变体实现 **最大池化**作为max运算，采用 $\mathcal{T}_{relu}$ 的变体（N元max）实现

5️⃣ 针对GNN的编译扩展

1. **前提**: 编译过程基于一个**给定的、固定的图拓扑** $G = (V, E)$
2. **聚合层**: **Sum/Mean聚合**作为线性操作，应用 $\mathcal{T}_{linear}$ 的稀疏变体实现 **Max聚合**作为max运算，应用 $\mathcal{T}_{relu}$ 的变体实现
3. **更新层**: GNN的更新层（MLP）则直接复用上述FFN的编译方法（$\mathcal{T}_{linear}$ 和 $\mathcal{T}_{relu}$ 的交替复合） 整个消息传递过程可被编译为一个在图上运行的**图变换器 (Graph Transducer)**

3.4. 形式化分析的应用

3.4.1. 基础验证与功能等价性

✅ **定理WFA-2：基于自动机的属性验证定理**

1. **陈述**: 设 F 为一个神经网络，其计算行为可由一个在半环 K 上的加权自动机 A_F 精确表示 设 ϕ 为一个待验证的性质，其所有**反例**构成的集合可以被一个在同一半环 K 上的加权自动机 $A_{\neg\phi}$ 所识别
   - 通过自动机的**复合(Composition)或等价的乘积构造(Product Construction)**，可得到一个新的自动机 $A_{cex} = A_F \circ A_{\neg\phi}$ 此自动机 A_{cex} 精确地描述了所有使网络 F 违反性质 ϕ 的输入集合
   - 因此，可得出以下确定性结论：
     - 若自动机 A_{cex} 所接受的语言为空（即 $L(A_{cex}) = \emptyset$），则网络 F **恒满足**性质 ϕ
     - 若 A_{cex} 所接受的语言非空，则性质 ϕ **不成立**，且 A_{cex} 本身即为所有反例输入的集合
2. **意义**: 它通过将“网络是否满足性质”这一复杂问题，转化为“某个自动机是否为空”这一经典可解问题，为神经网络的形式化验证提供了一个完备、无误的“终极裁判”

✅ **定理WFA-3：基于自动机的模型功能等价性判定定理 (修正版)**

1. **陈述**: 两个神经网络 F_1 和 F_2 （包括FFN、CNN及在**给定图上展开的GNN**）功能等价，即 $F_1(\mathbf{x}) \equiv F_2(\mathbf{x})$ 对所有输入 $\mathbf{x}$ 成立，当且仅当它们对应的加权自动机 A_{F_1} 和 A_{F_2} 等价 由于这些前馈网络生成的自动机是**无环的 (acyclic)**，该等价性问题在热带半环 $(\mathbb{R} \cup \{-\infty\}, \max, +)$ 上理论上是**可判定的**
2. **意义**: 将“两个神经网络功能是否完全一样”这个无法直接回答的黑盒问题，首次转化为一个有确定性答案（理论上可判定）的白盒数学问题

✅ **定理WFA-4：基于自动机的模型差异区域形式化表征定理**

1. **陈述**: 设有网络 F_1, F_2 及其编译后的自动机 A_{F_1}, A_{F_2}，给定差异阈值 $\epsilon \geq 0$ 存在一个确定性算法，能构造出一个新的自动机 $A_{D(\epsilon)}$，其精确定义了使两网络输出差异大于 ϵ 的所有输入构成的几何区域 $D(\epsilon)$：

$$L(A_{D(\epsilon)}) \equiv D(\epsilon) \triangleq \{\mathbf{x} \in \mathbb{R}^n \mid |F_1(\mathbf{x}) - F_2(\mathbf{x})| > \epsilon\}$$

2. **意义**: 它将模型比较从“它们是否相同”的简单判断，提升为对“它们在哪些输入上、以何种方式不同”的完整几何刻画

3.4.2. 函数几何的精确解剖

✅ **定理WFA-5： 基于自动机的输出值域精确分析定理**

1. **陈述**: 对于由神经网络 F 编译而来的无环加权自动机 A_F，其所计算的函数可表示为 $F(\mathbf{x}) = \max_{\pi \in \Pi}(\mathbf{w}_\pi^T \mathbf{x} + b_\pi)$，其中 Π 是自动机编码的所有线性区域的集合 给定输入域 $\mathcal{D}_{in}$，网络输出值域的边界由以下优化问题确定：
   - **上界 (Maximum Value)**:
   $$\max_{\mathbf{x} \in \mathcal{D}_{in}} F(\mathbf{x}) = \max_{\mathbf{x} \in \mathcal{D}_{in}} \left(\max_{\pi \in \Pi} (\mathbf{w}_\pi^T \mathbf{x} + b_\pi) \right)$$
   - **下界 (Minimum Value)**:
   $$\min_{\mathbf{x} \in \mathcal{D}_{in}} F(\mathbf{x}) = \min_{\mathbf{x} \in \mathcal{D}_{in}} \left(\max_{\pi \in \Pi} (\mathbf{w}_\pi^T \mathbf{x} + b_\pi) \right)$$
2. **意义**: 它将神经网络的函数彻底分解，使得计算其输出范围不再需要任何猜测或区间放大，而是能像解方程一样得到一个精确无误的最终答案

✅ **定理WFA-6：基于自动机的全局稳定性界定与Lipschitz常量精确计算定理**

1. **陈述**: 设神经网络函数 F 已被编译为加权自动机 A_F 该自动机的结构编码了函数所有线性区域 π 的梯度向量 $\mathbf{w}_\pi$ 存在一个确定性算法，能从 A_F 中提取出完整的梯度向量集合 $W_F = \{\mathbf{w}_\pi\}$ 因此，函数 F 在整个输入空间 $\mathbb{R}^n$ 上的全局 **Lipschitz 常数** L，对于任意 p-范数，均可被**精确计算**:
   $$L_p(F) = \max_{\mathbf{w} \in W_F} \|\mathbf{w}\|_p$$
2. **意义**: 它将衡量模型全局鲁棒性的黄金标准——Lipschitz常数，从一个过去只能靠放大来“估计”的模糊上界，变成了一个可以被直接“算出”的精确数值

✅ **定理WFA-7：基于自动机的决策边界精确几何表征定理**

1. **陈述**: 设神经网络任意两类 i, j 间的决策边界由函数 $g(\mathbf{x}) = F_i(\mathbf{x}) - F_j(\mathbf{x})$ 的零点集 $DB_{ij} = \{\mathbf{x} \in \mathbb{R}^n | g(\mathbf{x}) = 0\}$ 定义，且 $g(\mathbf{x})$ 已被编译为等价的加权自动机 A_g 那么：
   - 该决策边界 DB_{ij} 的几何本质是一个**有限多面复形 (Finite Polyhedral Complex)**
   - 一个能完全定义此复形的**有限几何约束集**（即所有超平面方程 $\mathbf{w}_\pi^T \mathbf{x} + b_\pi = 0$ 与半空间不等式 $\mathbf{w}_\pi^T \mathbf{x} + b_\pi \leq 0$）可从 A_g 中被**确定性地、完整地提取**
2. **意义**: 它实现了神经网络决策边界的终极“白盒化”，将这个模型内部隐式的、高维的几何结构，完整地“导出”为一个可被计算机直接分析和推理的精确数学对象

✅ **定理WFA-8：基于自动机的线性区域划分显式构造定理**

1. **陈述**: 设神经网络函数 $F(\mathbf{x}) = \max_{\pi \in \Pi}(\mathbf{w}_\pi^T \mathbf{x} + b_\pi)$ 已被编译为等价的加权自动机 A_F
   - 函数 F 将输入空间 $\mathbb{R}^n$ 划分为一个有限的凸多面体区域集合 $\{R_\pi\}_{\pi \in \Pi}$，在每个区域 R_π 内 F 的行为等同于单个仿射函数
   - 存在一个确定性算法，能从 A_F 中为**每个区域** R_π 构造出其完整的线性不等式约束集 C_π:
   $$R_\pi = \{\mathbf{x} \in \mathbb{R}^n \mid (\mathbf{w}_\pi - \mathbf{w}_\rho)^T \mathbf{x} + (b_\pi - b_\rho) \geq 0, \quad \forall \rho \neq \pi\}$$

2. **意义**: 它将分析粒度从网络决策的"边界线"，深化到了构成其内部逻辑的每一块"线性区域"，并能给出每个区域精确的数学"身份证"(即其完整的代数定义)

3.4.3. 模型可解释性与因果推断

✅ **定理WFA-9：基于自动机的梯度计算定理 (Automaton-based Gradient Computation Theorem)**

1. **陈述**: 对于一个表示神经网络 F 的加权自动机 A_F，其结构**隐式地定义**了梯度函数 $\nabla_{\mathbf{x}} F(\mathbf{x})$ 存在一个构造性算法，可以从 A_F 派生出一个**梯度计算机制**（如加权变换器 $T_{\nabla F}$） 该机制重用 A_F 的路径选择逻辑：当路径 π^* 对输入 $\mathbf{x}$ 激活时（即 $F(\mathbf{x}) = \mathbf{w}_{\pi^*}^T \mathbf{x} + b_{\pi^*}$），该机制输出对应的梯度向量 $\mathbf{w}_{\pi^*}$
2. **意义**: 它为解释模型"为何如此决策"的梯度归因方法，提供了一个不再依赖任何近似、保证对任意输入都绝对精确的计算引擎

✅ **定理WFA-10：形式化逆向分析与特征可视化定理**

1. **陈述**：给定一个由加权变换器 (Weighted Transducer) $\mathcal{T}_F$ 表示的网络函数 F，以及一个由自动机 $A_{\mathcal{Y}_{target}}$ 所描述的目标输出区域 $\mathcal{Y}_{target}$ 所有能使网络输出 $F(\mathbf{x})$ 落在该区域内的输入 $\mathbf{x}$ 的集合（即原像集 $\{\mathbf{x} | F(\mathbf{x}) \in \mathcal{Y}_{target}\}$），可以被一个等价的自动机 A_{preimage} 精确表示 该自动机通过对目标区域自动机应用**逆变换器 (Inverse Transducer)** 来构造：

$$A_{\text{preimage}} \equiv \mathcal{T}_F^{-1}(A_{\mathcal{Y}_{target}})$$

2. **意义**：它为神经网络提供了一个强大的"逆向搜索引擎"，能够形式化地、无遗漏地找到所有能导致特定结果（无论是好的还是坏的）的输入模式

✅ **定理WFA-11：基于自动机的形式化因果干预与影响量化定理**

1. **陈述**: 设网络 F 及其因果干预后的版本 F_C 分别被编译为等价的加权自动机 A_F 和 A_{F_C}
   - **表示**: 影响函数 $g_C(\mathbf{x}) \triangleq F(\mathbf{x}) - F_C(\mathbf{x})$ 的行为由自动机**对** (A_F, A_{F_C}) 完整定义
   - **量化**: 在给定的输入域 $\mathcal{D}_{in}$ 上，最大因果影响 $I_{\max}(C)$ 是一个可被精确计算的确定值：

$$I_{\max}(C) \triangleq \sup_{\mathbf{x} \in \mathcal{D}_{in}} |F(\mathbf{x}) - F_C(\mathbf{x})|$$

     该值可通过从自动机对 (A_F, A_{F_C}) 中提取所有线性片段的参数，构建并求解一个等价的**混合整数线性规划 (MILP)** 问题来获得
2. **意义**：它将"一个神经元或特征的真实贡献"这个模糊的概念，转化成了一个可以通过求解优化问题而得到的、具有明确上限的精确数值

3.4.4. 模型元属性与结构分析

✅ **定理WFA-12：基于自动机的函数最小复杂度定理**

1. **陈述**: 对于一个由网络 F 计算的函数，其对应的加权自动机 A_F 可以被最小化为一个等价的自动机 $A_{F,min}$ 这个最小自动机的状态数量，定义了该函数在自动机模型下的**最小实现复杂度**，是其内在复杂性的一个理论下界
2. **意义**: 它能"蒸馏"出神经网络所学函数最核心、不可再简化的"计算内核"，并用一个具体的数字（最小状态数）来量化其真实的内在复杂度。

✅ **定理WFA-13：逐层复合构造定理 (Hierarchical Composition Theorem)**

1. **陈述**: 令 $\mathcal{T}_l$ 为第 l 层网络对应的自动机变换器（Transducer），则代表整个深度为 L 的网络的自动机 A_F ，可以表示为所有层变换器的顺序复合：

$$A_F = \mathcal{T}_L \circ \mathcal{T}_{L-1} \circ \cdots \circ \mathcal{T}_1(A_{in})$$

其中 A_{in} 是代表输入的平凡自动机，$\circ$ 是自动机/变换器的复合算子

2. **意义**: 它将整个网络的编译过程模块化，为在任意中间层"暂停"并精确分析网络内部特征的演化提供了理论依据。

✅ **定理WFA-14：基于自动机的等变性验证定理 (Automaton-based Equivariance Verification Theorem)**

1. **陈述**: 令 g 为一个作用于输入的几何变换（如平移、置换），g' 为其在输出空间对应的变换 若 g 和 g' 均可被自动机变换器 $\mathcal{T}_g$ 和 $\mathcal{T}_{g'}$ 表示，则网络 F 具备 (g, g')-等变性，当且仅当以下自动机等价关系成立：

$$A_F \circ \mathcal{T}_g \equiv \mathcal{T}_{g'} \circ A_F$$

2. **意义**: 它将证明网络几何性质这一可能因"符号爆炸"而无法计算的理论问题，转化为一个具体、可由算法高效判定的自动机等价性问题。

✅ **定理WFA-15：基于自动机的决策边界对权重的微分几何灵敏度定理**

1. **陈述**:设由权重 W 参数化的神经网络 F_W，其决策边界 $DB_{ij}(W)$ 由有限超平面集 $\{H_\pi(W)\}_{\pi\in\Pi_g}$ 定义，其中 $H_\pi(W) : (\mathbf{w}'_\pi(W))^T\mathbf{x} + b'_\pi(W) = 0$ 那么：
   - 描述该边界的自动机 $A_{g,W}$ 是一个**符号参数化的加权自动机**，其权重是关于网络权重 W 的有理多项式函数
   - 存在一个确定性算法，能从 $A_{g,W}$ 构造出一个**边界灵敏度变换器**，用以精确计算决策边界中**每一个**超平面 $H_\pi(W)$ 对**每一个**网络权重 W_k 的偏导数：

$$\frac{\partial H_\pi(W)}{\partial W_k} \triangleq \left(\frac{\partial \mathbf{w}'_\pi(W)}{\partial W_k}, \frac{\partial b'_\pi(W)}{\partial W_k} \right)$$

2. **意义**：它首次建立了连接模型参数空间（权重）与输入空间（决策边界）的精确微分桥梁，使我们能从根本上量化和分析模型结构对自身参数扰动的几何稳定性。

✅ **定理WFA-16：关键图结构扰动的形式化识别定理**

1. **陈述**: 设 F 为一个在图 G 上定义的GNN，并给定一个**候选扰动边集** $E_{cand} = \{e_1, e_2, \ldots\}$ 该框架保证存在一个确定性的算法，可以从 E_{cand} 中精确识别出对网络输出造成最大影响的"**最关键扰动**" e^*：

$$e^* = \arg\max_{e_i \in E_{cand}} \left(\sup_{\mathbf{x}\in\mathcal{D}_{in}} |F(\mathbf{x}, G \cup \{e_i\}) - F(\mathbf{x}, G)| \right)$$

此过程通过为每个候选扰动 e_i 构造并分析其真实的差分函数自动机来精确实现

2. **意义**：它将对GNN的分析从被动地"评估"单个结构扰动的影响，提升为主动地"寻优"，能保证从众多可能性中精确找出导致模型输出变化最大的那条"最关键的边"。

3.4.5. 模型的综合与修复

✅ **定理WFA-17 : 模型修复问题的形式化构建定理**

1. **陈述**: 设有一个由权重 W 参数化的网络 F_W 及一个由自动机 A_Φ 定义的目标功能属性 Φ 寻找最小权重扰动 ΔW 以修复该网络的问题：

$$\begin{aligned} \min_{\Delta W} \quad & \|\Delta W\|_p \\ \text{s.t.} \quad & A_{W+\Delta W} \models \Phi \end{aligned}$$

可以被确定性地**形式化构建 (formally constructed)** 其约束条件定义了一个关于变量 ΔW 的**半代数集 (semi-algebraic set)**，该问题在理论上**可判定 (decidable)**，但在计算上通常是**不可行 (intractable)** 的

2. **意义**：它将“修复”神经网络这一高度启发式的工程任务，首次提升到了一个具有精确数学定义和理论可解性保证的约束优化问题层面，为该问题的复杂性分析和未来近似算法的设计奠定了基石。

✅ **定理WFA-18 : 模型结构修复问题的形式化构建定理**

1. **陈述**: 设网络 F (由自动机 A_F 表示) 违反了性质 Φ。令 $\mathcal{O} = \{o_1, o_2, \ldots\}$ 为原子结构编辑算子集。寻找最短编辑序列 $S \in \mathcal{O}^*$ 以修复网络的问题，可形式化构建为：

$$\begin{aligned} \min_{S \in \mathcal{O}^*} \quad & \text{length}(S) \\ \text{s.t.} \quad & A_{S(F)} \models \Phi \end{aligned}$$

2. **意义**：它将模型修复的范式从“调参”扩展到“改结构”，为“如何通过最少的增删组件来修复网络缺陷”这一根本问题提供了精确的数学定义