Exercise takeaways

April 21, 2026

From approximate condensation to approximate weak natural latent (Exercises 1 and 2). Given an ε-approximate condensation (Y_A) of observables $X_1, \ldots, X_n$, fix any subset $A \subseteq I$ and construct $\Lambda = (Y_B : B \cap A \neq \varnothing, |B| > 1)$, the collection of all latent variables contributing to more than one observable in A. Exercise 1 shows that Λ satisfies approximate mediation over X_A: the latent variable model condition implies that each X_i is a deterministic function of the latents above it, which bounds the residual dependence among observables after conditioning on Λ. Exercise 2 shows that Λ satisfies weak redundancy: each Y_B contributing to Λ is approximately recoverable from all-but-one observable in A, since the condensation score bounds how much information can fail to be captured by any large subset of observables. Together, these give an approximate weak natural latent from any condensation.

From approximate condensation to approximate strong natural latent (Exercise 3). If additionally there is no lower-level structure within A — formally, $Y_{\{i\}} \approx$ const for all singletons $\{i\} \subseteq A$, meaning the only shared information among observables in A flows through Λ rather than through pair-level latents — then weak redundancy upgrades to strong redundancy: Λ is approximately recoverable from any single observable in A, not just from all-but-one. This gives a full approximate natural latent.

From natural latent to condensation, and the agreement theorem (Exercises 4 and 5). Exercise 4 runs the construction in reverse: given any ε-approximate natural latent Λ over $X_1, \ldots, X_n$, setting $Y_I = \Lambda$ and all other $Y_A =$ const produces a 2ε-approximate condensation (error doubles because both mediation and redundancy each contribute ε). Exercise 5 uses a hierarchical structure: if observables admit a chain of natural latents $\Lambda^0 \supseteq \Lambda^1 \supseteq \cdots \supseteq \Lambda^n$ at successive levels of coarseness, the corresponding LVM forms a perfect (or approximate) condensation. The key point is that the Bayes net structure of the hierarchy — where each Λ^k_B is a parent of $\Lambda^{k-1}_{B'}$ whenever $B \supset B'$ — is precisely the structure required by Theorem 5.10 of the condensation paper for perfect condensation, and the strong redundancy condition at each level ensures each Y_A is recoverable from the upward cone $Y_{\supseteq A}$.

Why hierarchical natural latents form a perfect condensation. The condensation paper's Theorem 5.10 gives two necessary and sufficient conditions for a perfect condensation: (i) every Y_A is a deterministic function of any single observable X_i with $i \in A$, and (ii) the latent variables satisfy a Bayes net structure where Y_B is a parent of $Y_{B'}$ if and only if $B \supset B'$. A hierarchy of natural latents satisfies both. Condition (i) is just strong redundancy: $H(\Lambda^k_A \mid X_i) \approx 0$ for all $i \in A$ is exactly what it means for Λ^k_A to be a natural latent over X_A. Condition (ii) is satisfied by construction, since the hierarchy is indexed by nested subsets. The conditioned score $\chi(A)$ then collapses to $H(X_A)$ because the chain rule expansion telescopes: $H(\Lambda^0_A \mid \Lambda^1_A, \ldots) + H(\Lambda^1_A \mid \Lambda^2_A, \ldots) + \cdots$ collapses to $H(X_A)$ once strong redundancy at each level is used to show each conditional entropy term contributes negligibly beyond the bottom level.