

Programmer's Apprentice Preliminary CS379C Project Decomposition Plan

Tom Dean, Stanford, April 27, 2021

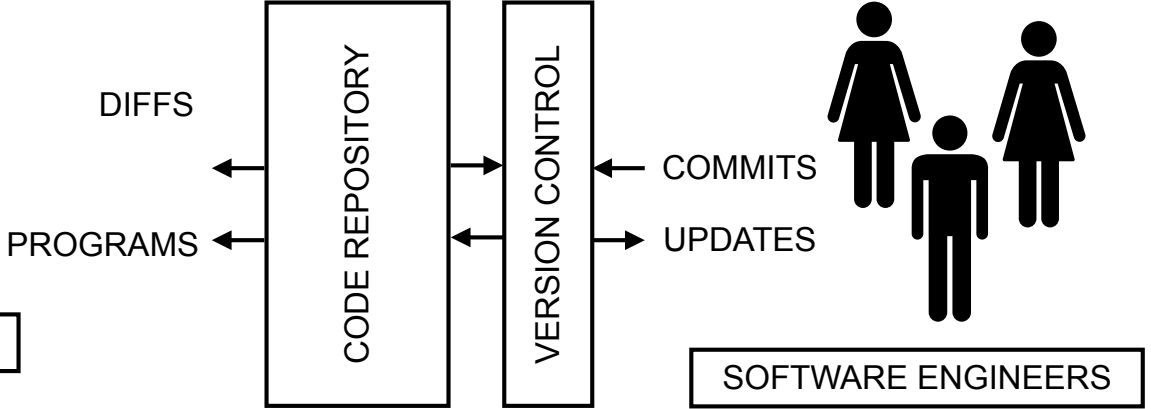
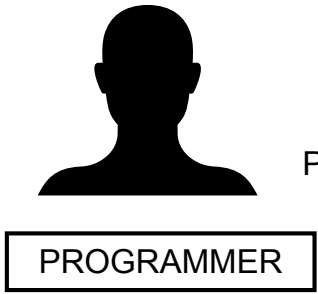
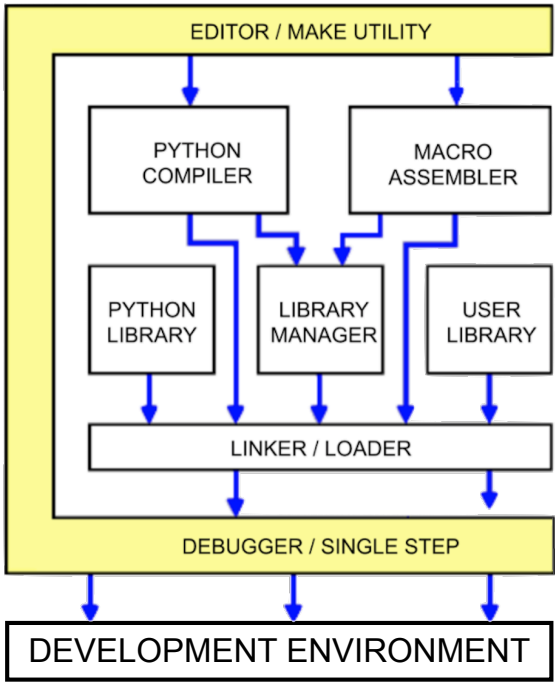


Acknowledgements: Francis Eugene Lewis, Yash Savani, Chaofei Fan



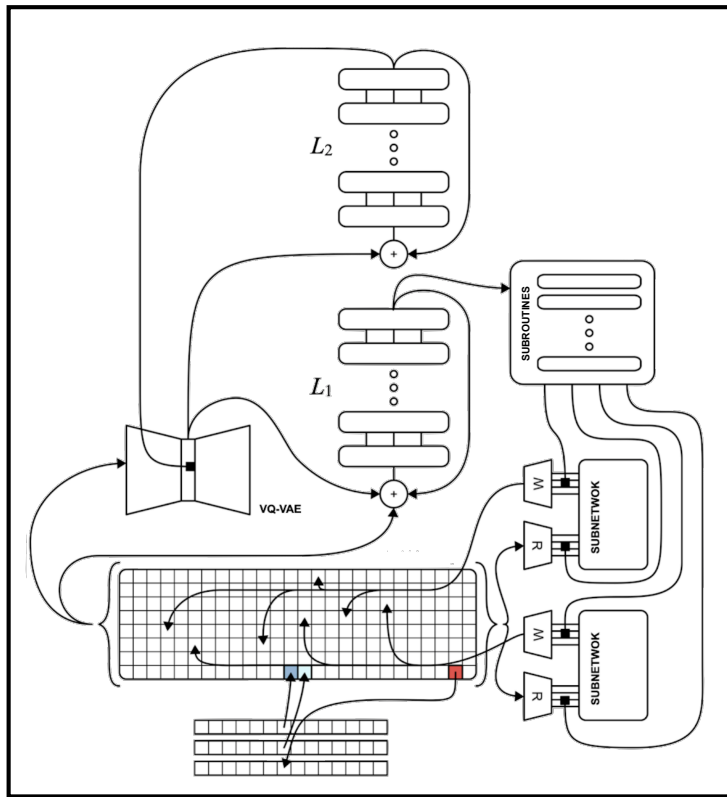
The programmer and assistant use gestures, including pointing and highlighting, to refer to entities in code listings – for example, variables, constants, type declarations, and syntactically complete expressions. Both use pointing to communicate relationships between entities corresponding to vertices in abstract syntax trees. They can invoke commands by pointing to the parameters of predefined templates that span the set of tasks expected of the apprentice. They are effectively grounded in separate domains, but they share a common core.

PROGRAMMER INTERACTS WITH THE APPRENTICE THROUGH A SHARED DEVELOPMENT ENVIRONMENT

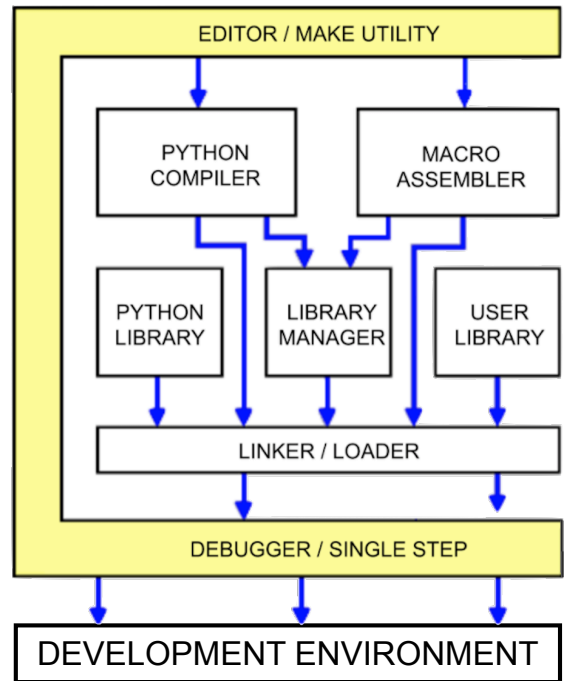


PROGRAMMER INTERACTS WITH THE WORLD OF SOFTWARE ENGINEERING THROUGH A CODE REPOSITORY

THE ASSISTANT INTERACTS DIRECTLY WITH THE PROGRAMMER THROUGH THE IDE ...



ASSISTANT NEURAL ARCHITECTURE



... AS WELL AS INDIRECTLY WITH THE WORLD OF CODE AND SOFTWARE ENGINEERING

HERE ARE TWO OF MANY POSSIBLE STARTING PLACES FOR DESIGNING THE PERCEPTUAL STACK

Perceiver: General Perception with Iterative Attention

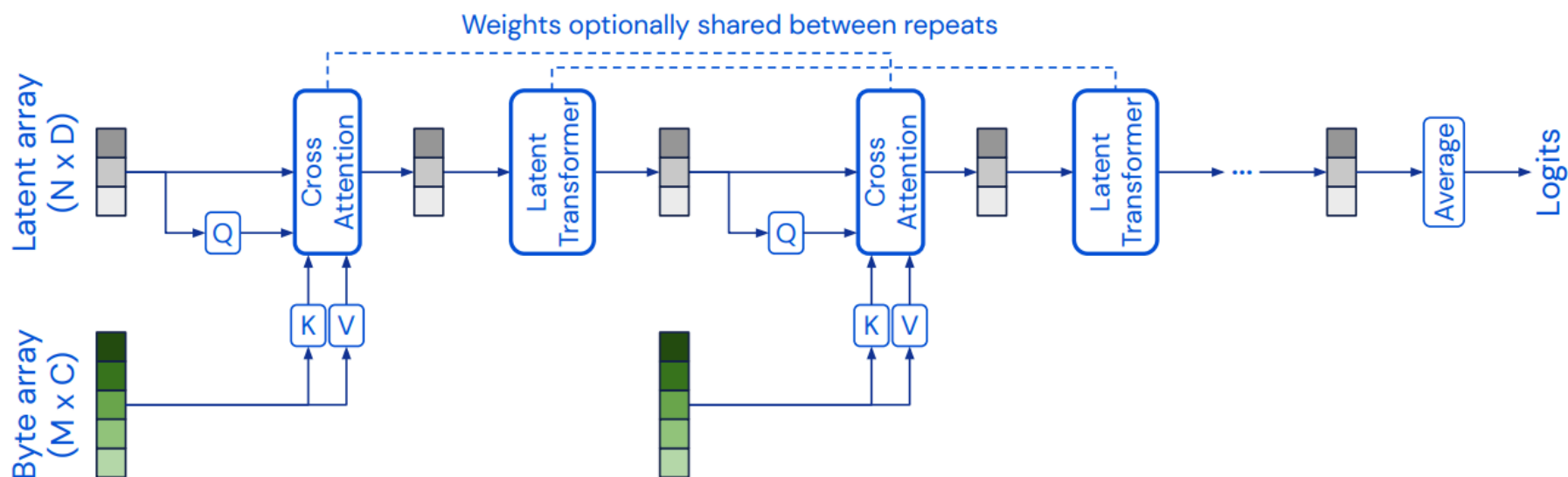


Figure 1. The Perceiver is an architecture based on attentional principles that scales to high-dimensional inputs such as images, videos, audio, point-clouds (and multimodal combinations) without making any domain-specific assumptions. The Perceiver uses a cross-attention module to project an input high-dimensional byte array to a fixed-dimensional latent bottleneck ($M \gg N$) before processing it using a stack of transformers in the low-d latent space. The Perceiver iteratively attends to the input byte array by alternating cross-attention and latent transformer blocks.

Tree Transformer: Integrating Tree Structures into Self-Attention

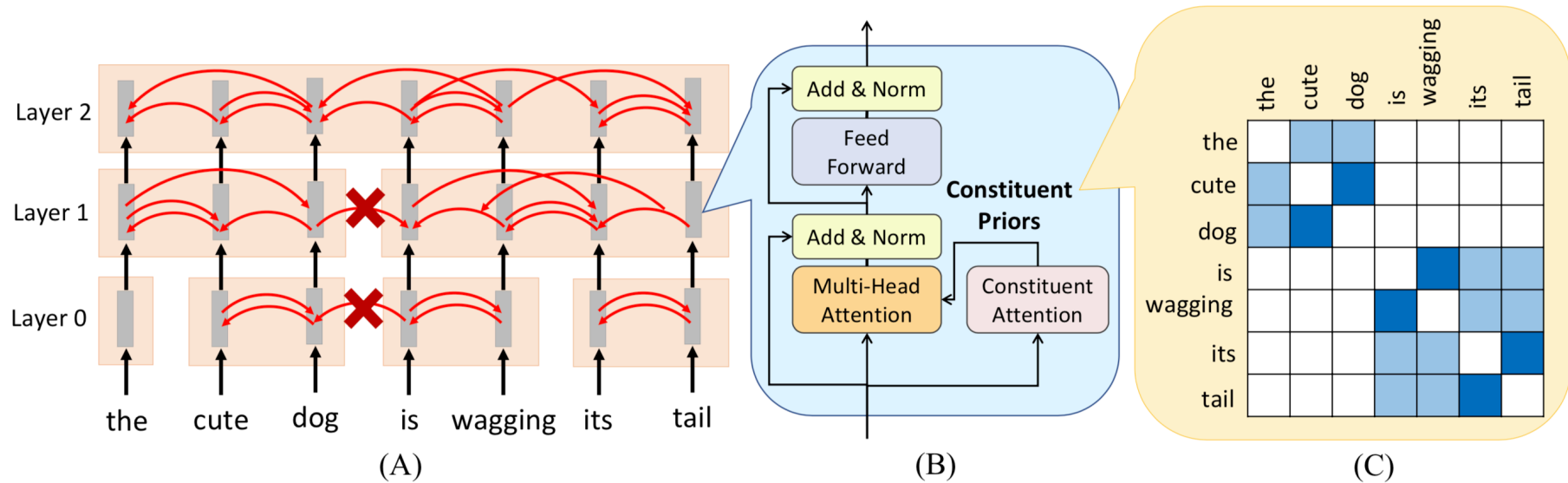


Figure 1: (A) A 3-layer Tree Transformer, where the blocks are constituents induced from the input sentence. The two neighboring constituents may merge together in the next layer, so the sizes of constituents gradually grow from layer to layer. The red arrows indicate the self-attention. (B) The building blocks of Tree Transformer. (C) Constituent prior C for the layer 1.

Complementary Learning Theory, One Shot Learning and Contrastive Predictive Coding

See: https://web.stanford.edu/class/cs379c/class_messages_listing/resources/Tom_Dean_One_Shot_Learning_Predictive_Coding_04_24_21.pdf