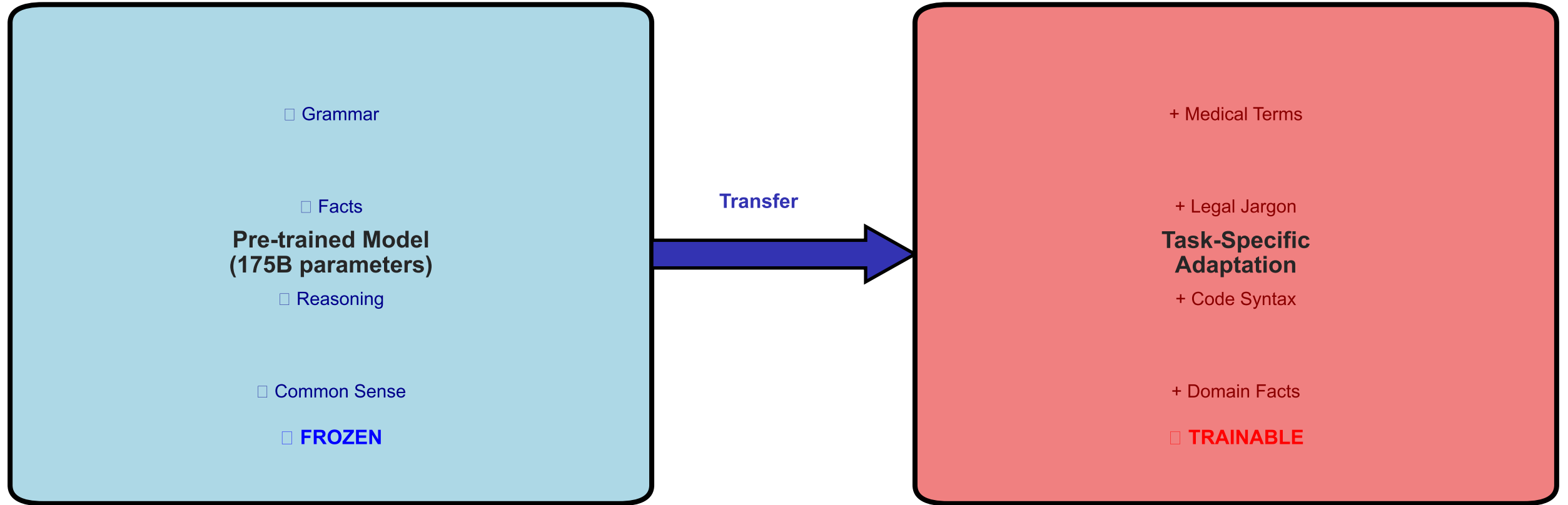


# Transfer Learning: Reuse General Knowledge, Adapt for Specifics



Traditional Approach:  
Train from scratch  
175B params, \$5M+



**Transfer Learning:**  
Freeze base, adapt top  
0.1-10% params, 100 – 5K