

Google TensorFlowがいろんなプロセッサに対応できるメカニズム

@Vengineer

GoogleのAI用コンパイラ TensorFlow XLAの基礎知識

● CPUは64ビットのx86とARMまで対応

人工知能を利用する際に、学習時は一般に強力なGPU (CUDA)が利用可能です。

判定時はGPUが利用できない組み込みシステムでは、CPU上に推論エンジンを実装することになります。TensorFlow XLAのAOTコンパイラは、TensorFlowで学習したモデルをCPU上で実行できる推論エンジンに変換できます注1(前頁下)。主な特徴を示します。

- (1) CPUはX86-64とARM64が正式にサポートされています。
- (2) CPUの実行コードを生成するにはLLVMを利用しています。

AOTコンパイラ部のソースコードを解析すると、X86-64とARM64以外に、ARM (32ビット)やPowerPCにも対応しているようです。

● 今のところ唯一! 新デバイスに対応可能なAI用コンパイラ

TensorFlow XLAは現在進行形で進化中です。AOTコンパイラ(現状CPUのみに対応)だけでなく、JITコンパイラもあり、図1に示すように、最新のTensorFlow XLAでは、今後出てくるであろう新しいハードウェアに対応することも可能になっています。

本稿執筆時点では、新しいハードウェアに対応するコードをリリースしているのは、TensorFlowだけです(本誌が皆さんのお手元に届くまでには、他のフレームワークでも新しいハードウェアに対応したコードをリリースする可能性は高いと思っています)。

TensorFlow XLA コンパイラの AI用新デバイスへの対応

● CPU/GPU以外に対応することを考えたAIフレームワークはあまりない

一般的には、新しいハードウェアを開発して市場に

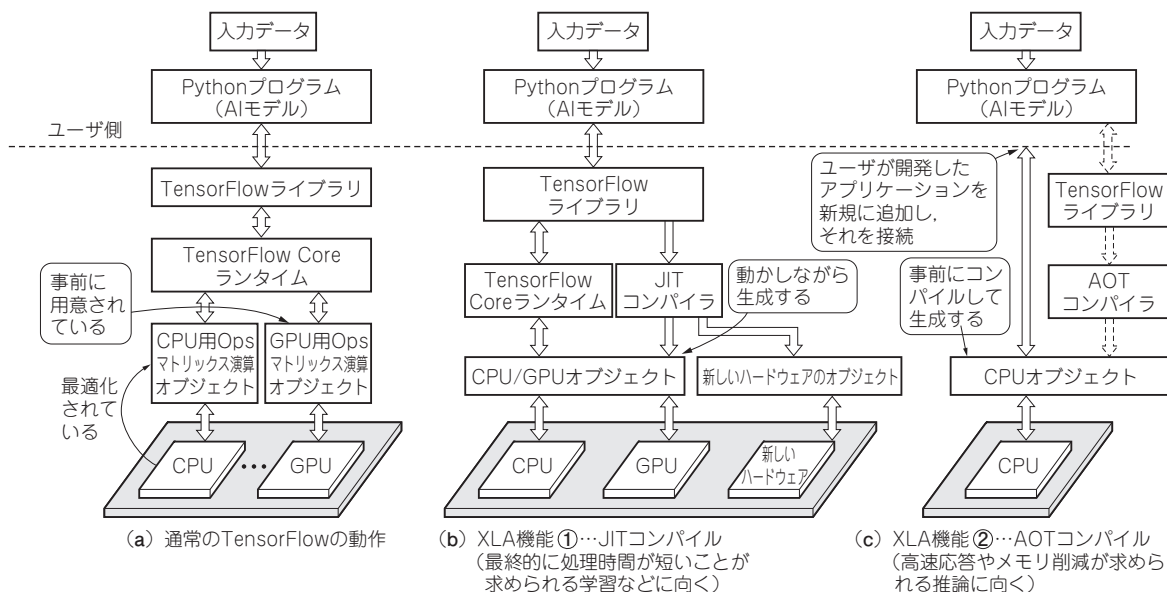


図1 GoogleのTensorFlow XLAコンパイラは現状唯一CPU/GPU以外のAI用新デバイスに対応できるようになっている