

機械学習を高効率に並列化する

川勝 孝也

木下 僚

高須 淳宏

安達 淳

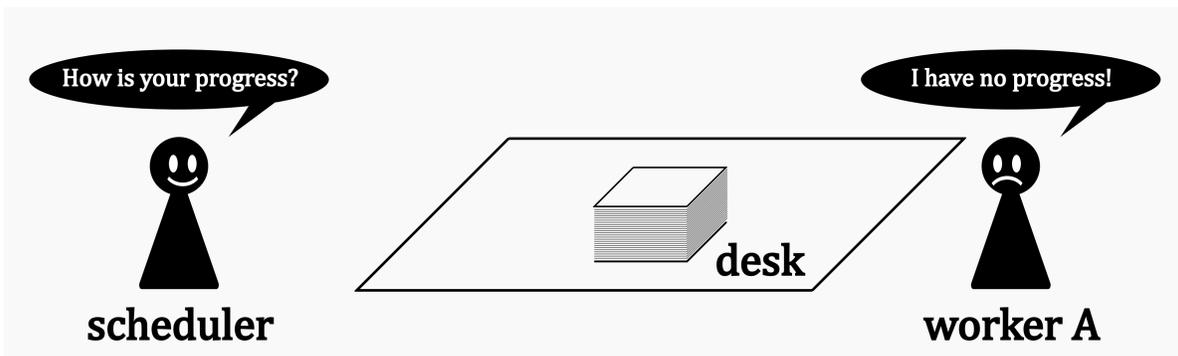
どんな研究？

人工知能に知識を与える機械学習は膨大な訓練データを大量のコアを持つ計算機で高速に処理します。**コア間で負荷のバランスが崩れると計算機が速度が低下する**のでコアの負荷をスケジューラで監視しますが、それ自体も速度が低下する原因になります。本研究では、勾配法を基礎とする機械学習全般を対象に、分割統治法を用いた並列処理により**バランス修正の頻度を抑制するADCA**を提案し、並列計算機の性能を最大限に引き出します。

何がわかる？

FIFOスケジューラが機械学習の並列処理の低速化を招くメカニズムを待ち行列理論で検証し、細粒度の並列化で低速化が顕著になることを確認しました。**分割統治法とワークスティーリングの併用**が速度低下の有効な対策となることを示し、機械学習フレームワークの設計指針を提案しました。並列処理を効率化することで人工知能に膨大な訓練データを学習させ、**賢い意思決定システム**を実現できるようになります。

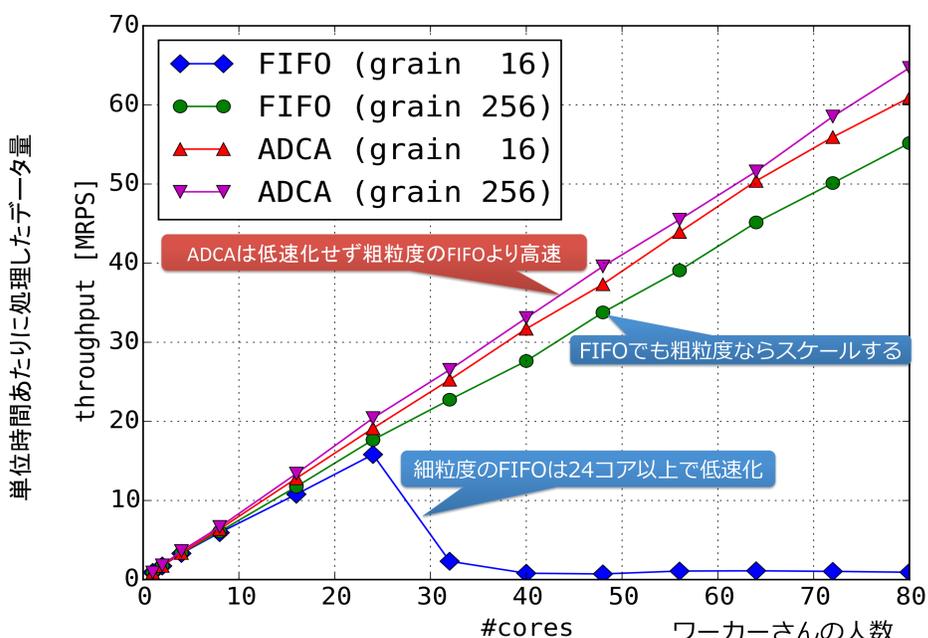
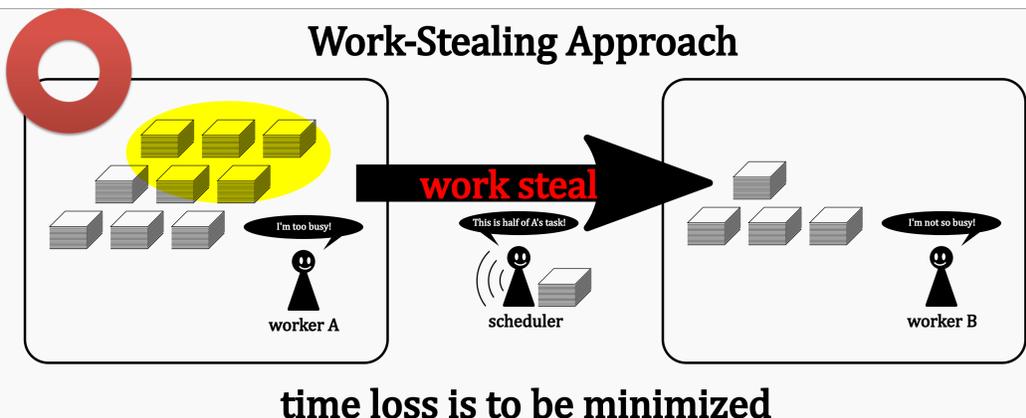
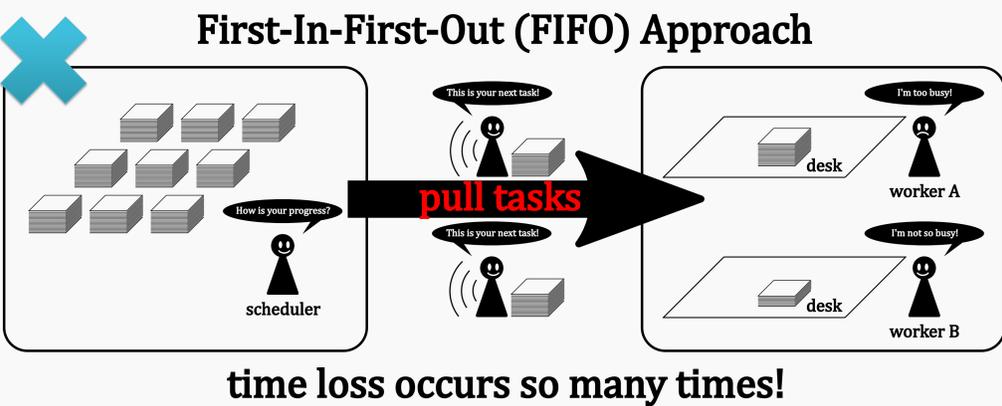
状況設定



並列計算機は、数十から数千のコアをワーカーとし、仕事をワーカーに分配して、手分けして処理します。**ワーカー全員が同じタイミングで仕事を終える**のが理想です。なぜなら、先に仕事を終えたワーカーが暇を持て余して、他のワーカーが忙しくしていると**計算機の全力が発揮できない**からです。ワーカーが忙しく仕事しているか監視する人を**スケジューラ**と呼びます。

研究内容

ワーカーさんたちは、仕事の能力に差があったり、たまたま仕事が簡単だったりして、仕事が完了するタイミングは揃いません。スケジューラさんは、最初から同じ量の仕事を与えるのを諦め、**仕事を小さなタスクに分けて**、少しずつ与えることにしました。



これを**FIFO**と言います。ワーカーさんたちは仕事を終えるたびにスケジューラさんから次の仕事を受け取りますが、受け取るまでに少し待ち時間が発生します。また、スケジューラさんはひとりしかいません。なので、ワーカーさんたちがいっせいに仕事をもらいにスケジューラさんに詰めかけると、スケジューラさんがパンクして仕事が進まなくなってしまいます。

ADCAは、最初から大きな仕事の塊をワーカーさんたちに分配してワーカーさんたちが互いに仕事を融通するワークスティーリングを導入しました。ワーカーさんが暇になると、**忙しいワーカーさんの仕事の半分をどっさり持って行きます**。仕事を融通し合う回数は**FIFO**でスケジューラさんから仕事をもらう回数よりも少なくなり、その分ワーカーさんはみっちり仕事に専念できます。