

Research 25: Probabilistic Data Processing and Approximation

担当：杉浦 健人（名大）

Incremental Knowledge Base Construction Using DeepDive

Jaeho Shin, Sen Wu, Feiran Wang, Christopher De Sa,
Ce Zhang, Christopher Re (Stanford Univ. / Univ. of Wisconsin-Madison)

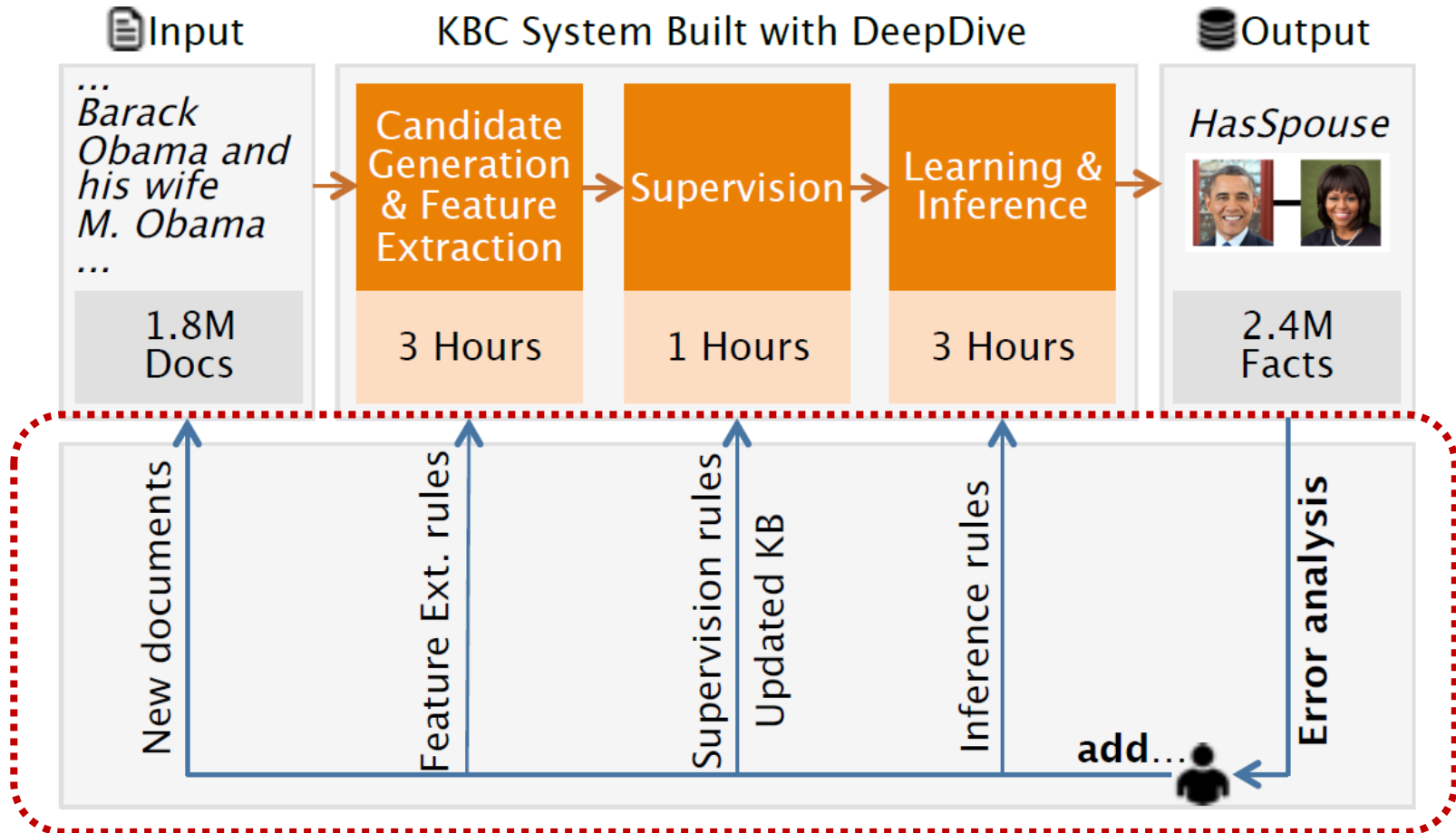
知識ベースの**インクリメンタルな更新**を提案

- ▶ 更新のたびに知識ベースを作りなおすより **22 倍速い**

アイデア

- ▶ **差分更新**の実行
- ▶ 推論における**信頼度計算の近似**
 - ▶ サンプルング
 - ▶ 有用でない推論ルールの削除

知識ベース構築の概要

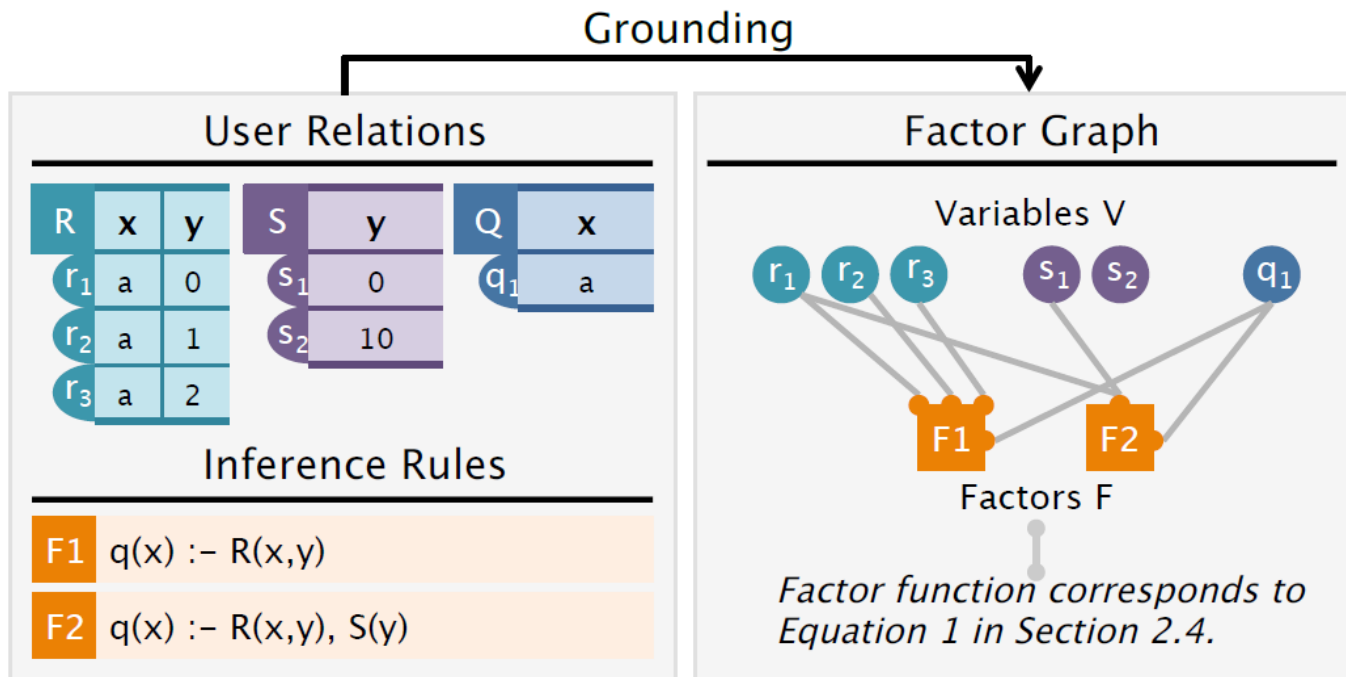


修正が行われたときにゼロから構築し直すのは無駄が多い

DeepDive における推論

推論は以下の 2 ステップ

1. Factor Graph による推論結果の生成
2. 推論結果の信頼度の計算



推論結果の更新

1. Factor Graph の更新

- ▶ データベースの View で実現可能

2. 推論結果の信頼度の更新

- ▶ 推論結果は非常に多いため全部更新するのは大変

1. サンプルングによる近似

2. Factor Graph の近似

- 推論に有用でないルールを削除

- ▶ データセットによって有効な近似手法は異なるため、有効な手法をシステムに自動で選択させる