# [6]Solving [9]Crossword [17]Puzzles

Anna Brezhneva, Rahul Mitra, Michael Stern

## Motivation

+ Thousands of daily solvers
+ Specific set of rules
+ Seemingly very human task

## Challenges

+ Huge state space
+ Semantic analysis is hard and domain-specific
+ Answers are more than just English words

## Problem Definition

+ fill in NYT crossword
+ More realistic: fill in smaller grid

## Approach

+ Use a CSP + Semantic Analysis
+ Variable is each word (not each letter)
+ Factors: word length (hard), word intersections (hard), clue (soft)
+ Backtrack search algorithm

## Data set

+ Past NYT data (1996-2012)
+ JSON Files representing the crossword puzzles
+ Over 400k clue/answer



## Backtrack

+ unassigned variable: MCV
+ Domain ordering: look for perfect match, otherwise compute score based on semantics and frequency
+ Lookahead: only look at one variable ahead

## Results

+ able to solve a **4x4** crossword with **100%** accuracy. puzzle was constructed using existing NYT clues
+ 15x15 NYT crossword takes a significant period of time and cpu runs at 100%

## Analysis

+ Looking forward: need to consider multiple word fills, right now the domain is only english words
+ In order to decrease run time for a full 15x15 puzzle, need more powerful ways to decrease domain