Learning Selection Strategies in Buchberger's Algorithm

Dylan Peifer, Michael Stillman, Daniel Halpern-Leistner

Cornell University

Buchberger's algorithm is

- a central tool for analyzing systems of polynomial equations
- the computational bottleneck in a wide variety of algorithms used in computer algebra software
- dependent for performance on human-designed decision heuristics at several key points in the algorithm

Buchberger's algorithm is

- a central tool for analyzing systems of polynomial equations
- the computational bottleneck in a wide variety of algorithms used in computer algebra software
- dependent for performance on human-designed decision heuristics at several key points in the algorithm

Idea: use reinforcement learning methods to train agents to make these decisions.

Main Contributions

1. Initiating the empirical study of Buchberger's algorithm from the perspective of machine learning.

2. Identifying a precise sub-domain of the problem that can serve as a useful benchmark for this and future research.

3. Training a simple model for pair selection which outperforms state-of-the art strategies by 20% to 40% in this domain.

Gröbner bases are special sets of polynomials that are useful in many applications, including

- computer vision
- cryptography
- biological networks and chemical reaction networks
- robotics
- statistics
- string theory
- signal and image processing
- integer programming
- coding theory
- splines
- ...

Question

Does the system of equations

$$\begin{cases} 0 = f_1(x, y) = x^3 + y^2 \\ 0 = f_2(x, y) = x^2 y - 1 \end{cases} \tag{1}$$

have an exact solution?

Question

Does the system of equations

$$\begin{cases} 0 = f_1(x, y) = x^3 + y^2 \\ 0 = f_2(x, y) = x^2 y - 1 \end{cases} \tag{1}$$

have an exact solution?

If there are polynomials a_1 and a_2 such that

$$h(x, y) = a_1(x, y)(x^3 + y^2) + a_2(x, y)(x^2 y - 1), \tag{2}$$

is the constant polynomial $h(x, y) = 1$, then there are no solutions.

Question

Does the system of equations

$$\begin{cases} 0 = f_1(x,y) = x^3 + y^2 \\ 0 = f_2(x,y) = x^2y - 1 \end{cases} \tag{1}$$

have an exact solution?

If there are polynomials a_1 and a_2 such that

$$h(x,y) = a_1(x,y)(x^3 + y^2) + a_2(x,y)(x^2y - 1), \tag{2}$$

is the constant polynomial $h(x,y) = 1$, then there are no solutions. If there are no solutions, then you can write 1 as a combination of $x^3 + y^2$ and $x^2y - 1$ by the weak Nullstellensatz (Hilbert, 1893).

Definition

The ideal generated by $f_1, \ldots, f_s$ is the set of all polynomials of the form

$$h = a_1 f_1 + \cdots + a_s f_s$$

where $a_1, \ldots, a_s$ are arbitrary polynomials.

Definition

Given a set of polynomials $F = \{f_1, \ldots, f_s\}$, the multivariate division algorithm takes any polynomial h and produces a remainder polynomial r, written $r = \text{reduce}(h, F)$, such that

$$h = q_1 f_1 + \cdots + q_s f_s + r$$

where the lead term of r is smaller than any lead term of the f_i.

Definition

A Gröbner basis G of a nonzero ideal I is a set of generators $\{g_1, g_2, \ldots, g_k\}$ of I such that the remainder $\text{reduce}(h, G)$ is guaranteed to be 0 if h is in I.

Theorem (Buchberger's Criterion, 1965)

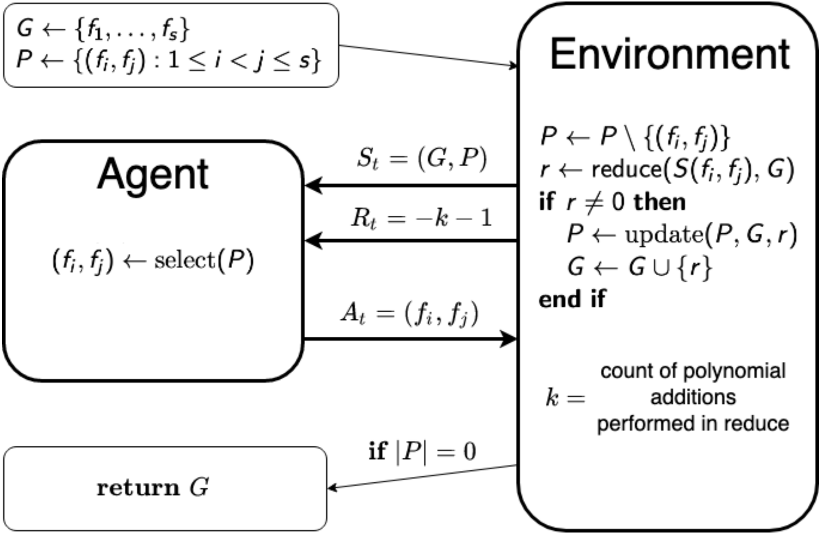
Suppose the set of polynomials $G = \{g_1, g_2, \ldots, g_k\}$ *generates the ideal* I*. If* $\text{reduce}(S(g_i, g_j), G) = 0$ *for all pairs* g_i, g_j*, where* $S(g_i, g_j)$ *is the S-polynomial of* g_i *and* g_j*, then* G *is a Gröbner basis of* I*.*

Example

In our previous example $F = \{x^3 + y^2, x^2y - 1\}$

$$\begin{aligned} r &= \text{reduce}(S(x^3 + y^2, x^2y - 1), F) \\ &= \text{reduce}(y(x^3 + y^2) - x(x^2y - 1), F) \\ &= y^3 + x \end{aligned}$$

so Buchberger's criterion is not satisfied.


$G \leftarrow \{f_1, \ldots, f_s\}$
$P \leftarrow \{(f_i, f_j) : 1 \leq i < j \leq s\}$
Environment
$P \leftarrow P \setminus \{(f_i, f_j)\}$
$r \leftarrow \text{reduce}(S(f_i, f_j), G)$
if $r \neq 0$ **then**
$P \leftarrow \text{update}(P, G, r)$
$G \leftarrow G \cup \{r\}$
end if
$k =$ count of polynomial additions performed in reduce
Agent
$(f_i, f_j) \leftarrow \text{select}(P)$
$S_t = (G, P)$
$R_t = -k - 1$
$A_t = (f_i, f_j)$
if $|P| = 0$
return G

Starting generators are binomials with no constant terms in 3 variables and a fixed maximum degree.

Example

$$\{x^3z + y^2, \quad x^2z^2 - xyz, \quad 5x^2y - 3z\}$$

Starting generators are binomials with no constant terms in 3 variables and a fixed maximum degree.

Example

$\{x^3z + y^2, \quad x^2z^2 - xyz, \quad 5x^2y - 3z\}$

- All new generators are also binomial.
- Some of the hardest known examples are binomial ideals.
- By adjusting the degree and number of initial generators, we can adjust the difficulty of the problem.

The state (G, P) is mapped to a $|P| \times 12$ matrix with each row given by the

$$(2 \text{ binomials})(2 \text{ terms})(3 \text{ variables}) = 12 \text{ exponents}$$

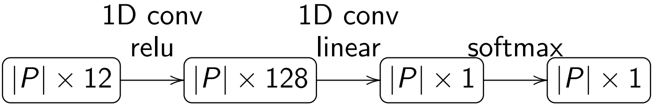
involved in each pair.

The state (G, P) is mapped to a $|P| \times 12$ matrix with each row given by the

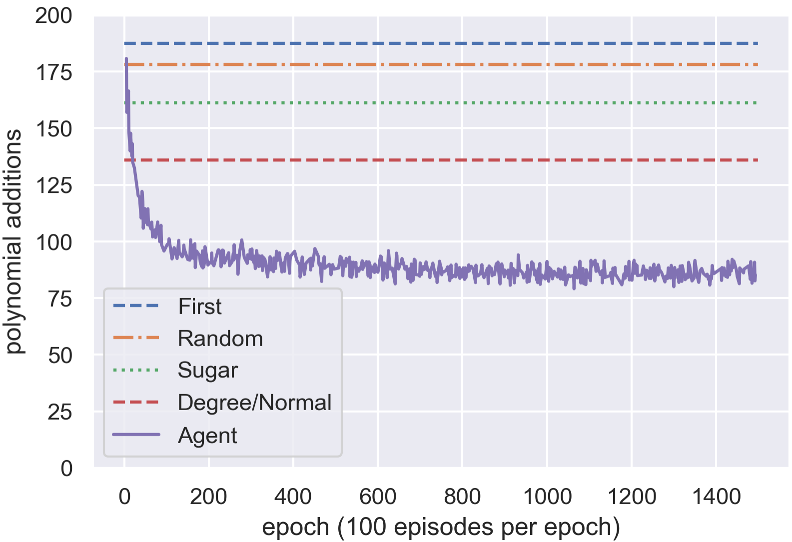
$$(2 \text{ binomials})(2 \text{ terms})(3 \text{ variables}) = 12 \text{ exponents}$$

involved in each pair.

This matrix is passed into a policy network



and a value model which computes the future return from following Degree selection.



200
175
150
125
100
75
50
25
0
polynomial additions
0
200
400
600
800
1000
1200
1400
epoch (100 episodes per epoch)
First
Random
Sugar
Degree/Normal
Agent

Summary

- Buchberger's algorithm is a central tool for analyzing systems of polynomial equations.
- Pair selection, a key choice in the algorithm, can be expressed as a reinforcement learning problem.
- In several distributions of random binomial ideals, our trained model outperformed state-of-the-art human-designed selection strategies by 20% to 40%.

Dylan Peifer djp282@cornell.edu
Michael Stillman mes15@cornell.edu
Daniel Halpern-Leistner daniel.hl@cornell.edu

https://github.com/dylanpeifer/deepgroebner