RESEARCH STATEMENT

Xinyu Sun

My research has been focused on *Computer-Assisted and Computer-Generated Mathematics*. The emergence of powerful mathematical computing environments, growing availability of correspondingly powerful computers and the pervasive presence of the Internet allow mathematicians to attack problems thought of being impossible just a decade ago. I use symbolic and numeric computation to search results, and whenever possible, prove them without human intervention. My study including the following areas:

1. Avoidable Patterns

A *word* is a finite sequence of letters over a predefined alphabet. A *square* is a word that is a concatenation of two identical words. A word is *square-free* if it contains no squares as factors. A binary word is a word over a two-letter alphabet, and a ternary word is over a three-letter alphabet. There are only six binary square-free words, but there are infinitely many ternary square-free words. The fact was proved (see in [LOT]) using the well-known Prouhet-Thue-Morse sequence. In fact, the number $s(n)$ of ternary square-free words of length *n* satisfies the following property: $S = \lim_{n\to\infty} s(n)^{1/n}$ exists, and *S* is called the "connective constant". Noonan and Zeilberger [NZ] proved that $S \le 1.30201064$, which is very close to the estimate of 1.302. However, finding the lower bound of S has posed as a programming challenge, as well as a theoretical one.

Definition 1 (*Brinkhuis triple*)**:** An *n*-Brinkhuis *k*-triple is three sets of words $B = \{B^0, B^1, B^2\}$, $B^i = \{w_j^i \mid 1 \le j \le k\}$, where w_j^i are square-free words of length *n*, such that for any square-free word $i_1 i_2 i_3, 0 \le i_1, i_2, i_3 \le 2$, and any $0 \le j_1, j_2, j_3 \le k$, the word $w_{j_1}^{i_1} w_{j_2}^{i_2} w_{j_3}^{i_3}$ of length *3n* is also square-free.

Grimm [GR] proved that the existence of an *n*-Brinkhuis *k*-triple implies $S \ge k^{1/(n-1)}$, and he subsequently proved that $S \ge 65^{1/40}$. I [SUN03] was able to prove the following:

Theorem 1: There exists a 43-Brinkhuis 110-triple, and thus $S \ge 110^{1/42}$.

Note: The result is best known yet. Unless there is a new way to find the lower bound or an algorithmic breakthrough in searching certain type of graphs, it appears we cannot improve the result by much.

2. Mathematical Games

Mathematical games, or combinatorial games, are two-player games with complete information (the players knows all the information about the game), no chance moves (no dice), a number of, usually finite many, positions, and the output is strictly win/lose or draw/draw. The player who makes the last move wins the game in the *normal play*, loses in the *misére play*. A game is *impartial* if every position is available to both players; otherwise it is *partizan*. We call a position a *P*-position if the player who makes the *previous* move will win, *N*-position if the *next* player will win.

Such games are interesting to us because they are quite unlike the traditional existential decision and optimization problems. Whereas in the existential decision problems area, there are only few problems whose complexity has not yet been determined, the complexity of the majority of combinatorial game is still unknown. A study of the precise nature of the complexity of those games enables us to attain a deeper understanding of the difficulties involved in certain new and old open game problems, which is a key to their solution. Research of these games can also lead us to new and interesting algorithmic challenges, in addition to the fun of playing games.

a. Chomp

Chomp is a two-player game that starts out with an *M* by *N* chocolate bar, in which the square on the top left corner is poisonous. A player must name a remaining square, and eat it together with all the squares below and/or to the right of it. Whoever eats the poisonous one (top-left) loses. The game can also be interpreted as two players alternately name a devisor of a given number *N*, which may not be multiples of previously named numbers. Whoever names 1 loses.

Some of the *P*-positions are easy to obtain: a position is a *P*-position if it has only two rows and the top row has one square more than the bottom; or if it has *n* columns with *n* squares in the first column and one square in the each of the other columns. Berlekamp, Conway and Guy [WW] gave a formula to find *P*-positions with other formations: if a position has *x* rows with *a* squares in the first row, *b* squares in the second and one square in the others, then it is a *P*-position if and only if

$$x = \begin{cases} \left\lfloor \dfrac{2a+b}{2} \right\rfloor & if \quad a+b \quad even; \\ \min\left\{ \left\lceil \dfrac{2a-b}{2} \right\rceil, \ \left\lceil \dfrac{3(a-b)}{2} \right\rceil \right\} & if \quad a+b \quad odd. \end{cases}$$

In [SUN02], I proved the following:

Theorem 2: If a position has *x* rows with *a* squares in the first row, *b* squares in the second, two squares in the third and at most one square in the rest of the rows, then it is a *P*-position if and only if

$$x = \begin{cases} 1 & \text{if} \quad a = 1; \\ 2 & \text{if} \quad a = b+1; \\ \left\lfloor \frac{2a+b}{2} \right\rfloor & \text{if} \quad a+b \quad \text{odd} \quad \text{and} \quad a \neq b+1; \\ \left\lfloor \frac{3a}{2} \right\rfloor + 1 & \text{if} \quad a = b; \\ \min\left\{ \left\lceil \frac{2a-b}{2} \right\rceil, \quad \frac{3(a-b)}{2} \right\} & \text{if} \quad a+b \quad \text{even} \quad \text{and} \quad a \neq b. \end{cases}$$

Another aspect of the research is to use computers to prove theorems. Computers have been used extensively to assist researchers to find a massive number of discrete results quickly and easily. However, using computers to prove theorems is still at its infancy. This is because computers in general do not understand logic and cannot calculate up to infinity. Using symbolic computing, Zeilberger [ZE] proved that with three-rowed Chomp, when the last rows have up to 115 squares, there are either finitely many *P*-positions, or the differences between the first two rows will eventually be constants.
I proved the following:

Theorem 3: For *P*-positions with k rows and $a_1, \ldots, a_k$ squares in the rows, $a_1 \leq \ldots \leq a_k$, when $a_1, \ldots, a_{k-2}$ are fixed, and either $k = 4$ and $a_{k-2} \leq 9$, or $k \leq 17$ and $a_{k-2} \leq 2$, either there exist only finitely many *P*-positions, or $a_k - a_{k-1}$ will eventually be periodic.

Theorem 4: If we denote g as the value of the Sprague-Grundy function of a Chomp position [$a_1, \ldots, a_k$], and if

$$\begin{cases} k = 2 \quad \text{and} \quad a_1 < 121, & \text{or} \\ k = 3 \quad \text{and} \quad a_1 \leq 3, a_1 + a_2 \leq 5, & \text{or} \\ k = 4 \quad \text{and} \quad a_1 = 1, a_1 + a_2 + a_3 \leq 3, & \end{cases}$$

then $g - a_k$ will be periodic when $a_1, \ldots, a_{k-1}$ are fixed and a_k large enough.

All these results are proved by computers without the need of human intervention, and each one demonstrates that with careful design, computers can be used to prove results involving infinity.

Byrnes [SB] later proved my conjecture that both of the results are true for all the *P*-positions with the corresponding top rows fixed.

Note: Theorem 2 is referenced by the recently issued 2nd edition of *Winning Ways for Your Mathematical Plays* by Berlekamp, Conway, and Guy.

b. Wythoff's Game

Wythoff's game [WYT] is an impartial game consisting of two piles of tokens. Players are to remove any number of tokens from a single pile, or the same number of tokens from both piles. The first player that cannot make a move loses. The game and its winning positions are well analyzed and explained in several papers.

A natural generalization of the game is the *N*-heap Wythoff's game, which consists of *N* piles of tokens, whose sizes are $A_1,\ldots,A_N$, $A_1 \le \ldots \le A_N$. A player can remove any number of tokens from a single pile, or remove ($a_1,\ldots,a_N$) tokens from all piles — a_i tokens from the ith pile, providing that $0 \le a_i \le A_i$, $\sum_{i=1}^{N} a_i > 0$, and $\bigoplus_{i=1}^{N} a_i = 0$, where $\oplus$ is the nim addition. Denote all the *P*-positions by $(A^1,\ldots,A^{N-2},A_n^{N-1},A_n^N)$, where $A^{N-2} \le A_n^{N-1} \le A_n^N$ and $A_{n-1}^{N-1} \le A_n^{N-1}$ for all $n \ge 0$. Two conjectures were preposed by Frænkel [FRA], which are also listed by Guy and Nowakowski [GN] as one of the "unsolved problems in combinatorial games", when $A^1,\ldots,A^{N-2}$ are fixed:

<u>Conjecture 1</u>: There exists an integer N_1 such that when $n \ge N_1$, $A_n^N = A_n^{N-1} + n$.

<u>Conjecture 2</u>: There exists integers N_2 and α_2 such that when $n \ge N_2$, $A_n^N = A_n^{N-1} = \lfloor n\phi \rfloor + \varepsilon_n + \alpha_2$ and $A_n^N = A_n^{N-1} + n$, where $\phi = (1+\sqrt{5})/2$, the golden section, and $-1 \le \varepsilon_n \le 1$.

Results have been speculated, but none of them was rigorously proved. I proved a sufficient condition for the conjectures for the three-heap Wythoff's game:

Denote $[m, A_n^m, B_n^m]$ as a *P*-position with *m* tokens in the first pile, $m + A_n^m$ second and $m + A_n^m + B_n^m$ third; N_2^m, α^m as the integers N_2, α_2 in the second conjecture when $A^1 = m$; $\alpha_n^m = A_n^m + m - \lfloor B_n^m \phi \rfloor$; $S_1^m(n) = \{i : 0 \le i \le B_n^m\} - \{B_i^m : i \ge 1\}$. Doron Zeilberger and I [SUN04] proved the following theorem:

Theorem 5: The following conditions imply both Conjecture 1 and 2 for a given *m*:
If there exists integers *N* and $n_1 \ge n_2 \ge N$ such that

- $A_N^m > \max\{b : \exists i < m \quad such \quad that \quad [i, m-i, b] \quad is \quad a \quad P-postition\}$
- $A_{n_2+3}^m + B_{n_2+3}^m \le A_{n_2}^m$;
- $B_{j+1}^m = B_j^m + 1$ for $n_2 \le j \le n_1$;
- $mex(\{B_i^m : i \le n_1\} \cup S_1^m(n_1)\} = B_{n+1}^m + 1$;
- $\max(\alpha_j^m : n_2 \le j \le n_1) - \min(\alpha_j^m : n_2 \le j \le n_1) \le 2$;
- $A_{n_2}^m > A_{N_2^i}^i, i < m$;

Furthermore if we denote $\alpha' = \lfloor (\max(\alpha_j^m : n_2 \le j \le n_1) - \min(\alpha_j^m : n_2 \le j \le n_1))/2 \rfloor$ and $\varepsilon_n^m = A_n^m + m - \lfloor B_n^m \phi \rfloor - \alpha'$, we assume:

- $\alpha^i - \alpha' \geq 4(m-i), 0 \leq i < m$;
- $\varepsilon_i^m \varepsilon_{i-1}^m \varepsilon_{i-2}^m = 0, n_2 < j < n_1$.

Applying the theorem above, we proved the following:

Theorem 6: Both of the conjectures are correct for three-heap Wythoff's game when the first heap has up to 10 tokens.

In addition to the theorems above, I [SUN05] defined and proved the following:

Definition 2 (*Wythoff's Sequence*)**:** We call a sequence of pairs of integers $\{(A_n, B_n)\}_{n \geq n_0}$ a Wythoff's sequence if there exist a finite set of integers T so that $A_n = mex(\{A_i, B_i : 0 \leq i < n\} \cup T)$, $B_n = A_n + n$ and $B_n \cap T = \phi$.

Definition 3 (*Special Wythoff's Sequence*)**:** A special Wythoff's sequence is a Wythoff's sequence such that there exist integers N and α so that when $n > N$, $A_n = \lfloor n\phi \rfloor + \alpha + \varepsilon_n$, where $\varepsilon_n \in \{0, \pm 1\}$.

Theorem 7: Every Wythoff's sequence is special. Therefore the two conjectures on the N-heap Wythoff's game are equivalent.

Theorem 8: Given a Wythoff's sequence $\{(A_n, B_n)\}_{n \geq n_0}$, there exists an integer c, so that $A_{A_n + c} = B_n - 1$ and $A_{B_n + c} = B_{A_n + c} + 1 = A_n + B_n + c$. Furthermore, $\alpha = -c$.

Following the ideas above, I plan to continue the study of mathematical games and avoidable patterns. They are endless open problems in the mathematical games, and each one of them demands different approaches. Even with the games that have been analyzed extensively, there are often rooms for improvement. For example, in spite of the result of Byrnes, finding a constructive winning strategy for Chomp still seems a long way from completion. For the N-heap Wythoff's game, it is desirable to eventually prove the two conjectures. Also, even if the conjectures are proved to be true, finding the winning strategies, especially the ones in polynomial time, still poses as a serious challenge. Moreover, we can ask the following question: are there any other multi-heap Wythoff's games whose winning positions involve irrational numbers other than the golden section?

Besides of the ternary square-free words, there are other avoidable patterns in disguise. For example *n*-dimensional random work is a set of word over an alphabet of *2n* letters, of which we trying to avoid factors such that they have equal number of letters *i* and *n*+*i*, for each $1 \leq i \leq n$. Finding the exact values of the upper bounds and lower bounds of the number of such words are still intriguing problems nowadays.

References:

[WW] E. R. Berlekamp, J.H. Conway, and R. K. Guy, *Winning Ways for Your*

Mathematical Plays, Academic Press, London, 1982. 598-599
[FRA] A. S. Frænkel, Complexity, appeal and challenges of combinatorial Games, to appear in *Theoret. Comp. Sci.*, special issue on Algorithmic Combinatorial Game Theory.
[GN] R. K. Guy and R. J. Nowakowski, Unsolved problems in combinatorial games, in: *More Games of No Chance*, Cambridge University Press, Cambridge, 2002, 457-473.
[LOT] M. Lothaire, *Combinatorics on words*, Addison-Wesley, 1983
[NZ] John Noonan and Doron Zeilberger, The Goulden-Jackson cluster method: extensions, applications, and implementations. *J. Difference Eq. Appl.* **5** (1999), 355-377.
[GR] Uwe Grimm, Improved bounds on the number of ternary square-free words, *J. of Integer Seq.* **4** (2001) 01.2.7
[SUN02] Xinyu Sun, Improvements on Chomp, *Integers*, **2** (2002)
[SUN03] Xinyu Sun, New Lower-Bound On The Number of Ternary Square-Free Words, *J. of Integer Seq*. **6** (2003), Article 03.3.2
[SUN04] Xinyu Sun and Doron Zeilberger, On Frænkel's *N*-Heap Wythoff's Conjectures, submitted
[SUN05] Xinyu Sun, Wythoff's Sequence and *N*-Heap Wythoff's Conjectures, submitted
[ZE] Doron Zeilberger, Three-Rowed CHOMP, *Adv. Appl. Math.* **26** (2001), 168-179
[WYT] W. A. Wythoff, A modification of the game of Nim, *Nieuw Arch. Wisk.* **7** (1907) 199-202.