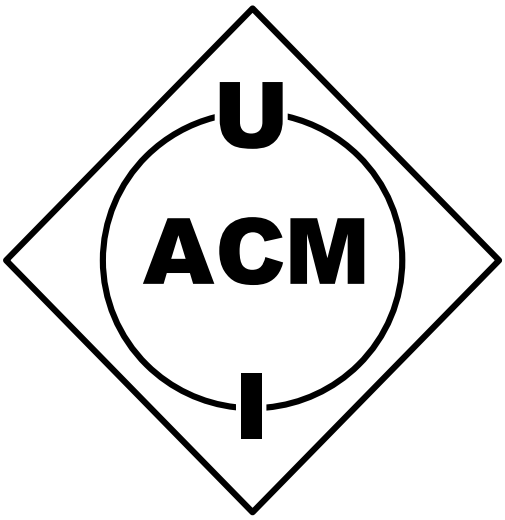
ACM@UCI Handbook

ICPC Templates



Ryan Yoshida

Contents

Part I
Introduction

1 About

1.1 Purpose

The ACM@UCI Handbook is a printable resource intended for use in contests such as ICPC. It provides data structure templates and code snippets for quick and reliable implementation. Priority is placed on data structures and algorithms which meet one or several of the following criteria: commonly seen in contests, difficult/tedious to remember/implement, and/or too niche to find in other competitive programming handbooks.

This handbook is not intended to be used as a teaching resource. Explanations of underlying theory and concepts for the provided data structures and algorithms will be provided only insofar as it is expedient to their implementation in contest. To make the best use of this handbook you should first have a strong understanding the purpose and theory behind every data structure and algorithm.

1.2 Organization

Part II
Data Structures

2 Graphs

2.1 Link-Cut Tree

2.2 Splay Tree

2.3 Union Find

3 Range Queries

3.1 Segment Tree

3.1.1 Point Update

3.1.2 Lazy Propagation

3.2 Square Root Decomposition

3.3 Sparse Table

3.4 Wavelet Tree

4 String Processing

4.1 Aho Corasick

Aho Corasick creats a finite automata on a prefix trie (4.4) to match occurences of a set of strings in a search string. Equivalently, it can be seen as an implementation of Knuth Morris Pratt (8.1) for a set of strings rather than one pattern string. It works the same as KMP by creating pointers to the longest proper suffix that is a prefix for each index of each string in the set. Construction of the Aho Corasic Trie is $O(M)$ where M is the sum of character counts of all pattern strings. Iteration through the search string is worst case $O(NK)$ where N is the length of the search string and K is the number of pattern strings. In most cases however, it is safe to assume that iteration through the search string is $O(N)$.

Implementation: The Aho Corasick Trie is encapsulated in a class. Due to the nature of the Aho Corasick automata construction algorithm, an `Aho` object must be constructed with a vector of pattern strings `dict`. To traverse the Aho Corasick Trie, a helper class `Aho::AhoAutomata` is provided. The `AhoAutomata` represents a state (node) in the Aho Corasick Trie. An iterator is provided for the `AhoAutomata` to iterate through the *ids* of all pattern strings in `dict` that are terminated at that node (the *id* of a pattern string is its index in `dict`). Note that the `Node` struct stores the child pointers as an `std::array` of length 256. This can be adjusted based on the problem.

```
#include <algorithm>

#include <vector>
#include <array>
#include <queue>
#include <string>

class Aho {
private:
    struct Node {
        Node *suffix;   // longest prefix that is suffix
```

```
        Node *output;   // longest word that is a suffix
        std::vector<int> wordids;   // ids of all words terminated at node
                                    // vector to handle duplicates
        std::array<Node *, 256> c;  // links to children of node

        Node()
        : suffix{nullptr}, output{nullptr}, wordids{std::vector<int>()} {
            std::fill(std::begin(c), std::end(c), nullptr);
        }
    };

private:
    Node *root;

public:
    class AhoAutomata {
    private:
        Aho::Node *root, *curr;
    public:
        AhoAutomata(Aho::Node *root) : root{root}, curr{root} { }

        // Advances to the next character, follows suffix link on failure
        void next(char c) {
            Aho::Node *tmp = curr;
            while (tmp->c[c] == nullptr && tmp != root) { tmp = tmp->suffix; }
            if (tmp->c[c] != nullptr) { curr = tmp->c[c]; }
            else { curr = root; }
        }

        class Iterator {
        private:
            Node *outputnode;
            int index;
        public:
            Iterator(Node *outputnode) : outputnode{outputnode}, index{0} {
                if (this->outputnode != nullptr
                        && this->outputnode->wordids.empty()) {
                    this->outputnode = outputnode->output;
                }
            }

            int operator*() { return outputnode->wordids[index]; }
            Iterator &operator++() {
                if (outputnode != nullptr) {
                    if (++index >= (int) outputnode->wordids.size()) {
                        outputnode = outputnode->output;
                        index = 0;
                    }
                }
                return *this;
            }

            bool operator==(const Iterator &other) {
                return outputnode == other.outputnode && index == other.index;
```

```
            }
            bool operator!=(const Iterator &other) {
                return outputnode != other.outputnode || index != other.index;
            }
        };

        AhoAutomata::Iterator begin() { return AhoAutomata::Iterator(curr); }
        AhoAutomata::Iterator end() { return AhoAutomata::Iterator(nullptr); }
    };

private:
    // Perform BFS to construct suffix and output links
    void computeLinks() {
        root->suffix = root;    // suffix link of root is itself

        std::queue<Node *> q;
        for (Node *next : root->c) if (next != nullptr) {
            next->suffix = root;    // edge case
            q.push(next);
        }

        while (!q.empty()) {
            Node *curr = q.front(); q.pop();
            for (int c = 0; c < 256; ++c) if (curr->c[c] != nullptr) {
                Node *suffixofcurr = curr->suffix;
                while (suffixofcurr->c[c] == nullptr && suffixofcurr != root)
                    suffixofcurr = suffixofcurr->suffix;

                if (suffixofcurr->c[c] != nullptr)
                    curr->c[c]->suffix = suffixofcurr->c[c];
                else
                    curr->c[c]->suffix = root;

                q.push(curr->c[c]);
            }
            if (!(curr->suffix->wordids.empty()))
                curr->output = curr->suffix;
            else
                curr->output = curr->suffix->output;
        }
    }

public:
    Aho(const std::vector<std::string> &dict) : root{new Node()} {
        for (int i = 0; i < static_cast<int>(dict.size()); ++i) {
            Node *curr = root;
            for (const char &c : dict[i]) {
                if (curr->c[c] == nullptr) { curr->c[c] = new Node(); }
                curr = curr->c[c];
            }
            curr->wordids.push_back(i);
        }

        computeLinks();
```

```
    }

    ~Aho() {
        std::queue<Node *> q;
        q.push(root);
        while (!q.empty()) {
            Node *curr = q.front(); q.pop();
            for (int c = 0; c < 256; ++c) if (curr->c[c] != nullptr) {
                q.push(curr->c[c]);
            }
            delete curr;
        }
    }

    AhoAutomata getAutomata() { return AhoAutomata(root); }

};
```

Example Usage: An example of how to use the Aho class is provided below.

```
// s is the search string, p is the pattern string set
void aho_example(const std::string &s, const std::vector<std::string> &p) {
    Aho aho(p);
    Aho::AhoAutomata automata = aho.getAutomata();

    for (int i = 0; i < (int) s.length(); ++i) {
        automata.next(s[i]);
        for (const int &id : automata) {
            std::cout << p[id] << "␣terminates␣at␣index␣" << i << std::endl;
        }
    }

    return 0;
}
```

4.2 EerTreE

4.3 Suffix Array

4.3.1 Longest Common Prefix

4.4 Trie

Part III
Algorithms

Computes an arithmetic sum constant time. The function was intentionally designed to have a similar function signature to Python's `range` function. Thus, `arithmetic_sum(start,end,step)` will yield the same result as `sum(range(start,end,step))` in Python, but it will do so in $O(1)$. For this function to work, $start < end$ and $step > 0$.

Important: `start` is inclusive, `end` is exclusive.

```
int arithmetic_sum(int start int end, int step=1) {
    int n = 1+(end-start-1)/step;
    if (n%2) return (start + ((n-1)>>1)*step)*n;
    else return (2*start + (n-1)*step)*(n>>1);
}
```

```
def arithmetic_sum(start: int, end: int, step: int=1):
    n = 1+(end-start-1)//step;
    if n%2:
        return (start + ((n-1)>>1)*step)*n
    else:
        return (2*start + (n-1)*step)*(n>>1)
```

7.2 Chinese Remainder Theorem

7.3 Division Over Modulo

Computes $\frac{a}{b} \mod mod$ by multiplying `a` to the inverse mod of `b`. Uses the power over modulo function described in 7.9. This function only works if `b` and `mod` are relatively coprime (i.e. $gcd(b, mod) = 1$). Can be used to compute inverse modulo by setting `a=1`. The time complexity of this algorithm is $O(\log_2 mod)$.

```
int divide_modulo(int a, int b, int mod) {
    return (a * (pow(b, mod-2, mod) % mod)) % mod;
}
```

```
def divide_modulo(a: int, b: int, mod: int):
    return (a * (pow(b, mod-2, mod) % mod)) % mod
```

7.4 Fast Fourier Transform

7.5 Geometric Sum

7.5.1 Finite Bounds

Returns $\sum_{k=start}^{end-1} a^k = \frac{a^{start}-a^{end}}{1-a}$.

```
double geometric_sum(int start, int end, double a) {
    return (std::pow(a,start) - std::pow(a,end))/(1.0-a);
}
```

```
def geometric_sum(start: int, end: int, a: float):
    return (pow(a,start) - pow(a,end))/(1-a)
```

7.5.2 Finite Start, Infinite End

Returns $\sum_{k=start}^{\infty} a^k = \frac{a^{start}}{1-a}$.

```
double geometric_sum(int start, double a) {
    return (std::pow(a,start))/(1.0-a);
}
```

```
def geometric_sum(start: int, a: float):
    return pow(a, start)/(1-a)
```

7.5.3 Infinite Scaled Sum

Returns $\sum_{k=0}^{\infty} ka^k = \frac{a}{(1-a)^2}$

```
double geometric_sum(double a) {
    return a/((1-a)*(1-a));
}
```

```
def geometric_sum(a: float):
    return a/((1-a)*(1-a))
```

7.6 Greatest Common Divisor

The algorithm for finding greatest common divisor is Euclid's algorithm. The time complexity of this algorithm is $O(\log_{10} max(a, b))$.

```
int gcd(int a, int b) { return b == 0 ? a : gcd(b, a % b); }
```

```
def gcd(a: int, b: int):
    return a if b == 0 else gcd(b, a%b)
```

7.7 Lowest Common Multiple

The algorithm for finding lowest common multiple uses Euclid's algorithm for greatest common divisor (7.6). The time complexity of this algorithm is $O(\log_{10} max(a, b))$.

```
int lcm(int a, int b) { return a * (b / gcd(a,b)); }
```

```
def lcm(a: int, b: int):
    return a * (b // gcd(a,b))
```

7.8 Modulo

A safe modulo for use with negative numbers. Use for languages like C++ when `a` can be negative.

```
int modulo(int a, int b) {
    const int res = a%b;
    return res >= 0 ? res : res + b;
}
```

7.9 Power Over Modulo

Returns $b^k \mod mod$. This implementation cannot handle $b < 0$ in C++. To handle this case, use the safe modulo described in section 7.8. Note that no Python implementation is provided because it is already implemented in the standard library function `pow(a,b[,mod])`. The time complexity of this algorithm is $O(\log_2(k))$.

```
int pow(int b, int k, int mod) {
    if (k == 0) return 1 % mod;
    else if (k == 1) return b % mod;
    else {
        int res = pow(b, k/2, mod);
```

```
        return (((res*res)%mod) * ((k%2 ? b : 1)%mod))%mod;
    }
}
```

8 Strings

8.1 Knuth Morris Pratt: String Matching