187. Repeated DNA Sequences

题目描述： **https://leetcode.com/problems/repeated-dna-sequences/**

> 给定一个基因序列，找到长度为10的重复序列。
> 例如：
>
> ```
> s = "AAAAACCCCCAAAAACCCCCAAAAAGGGTTT"
> ```
>
> 返回：
>
> ```
> ["AAAAACCCCC", "CCCCCAAAAA"]
> ```

解题思路：

将10bit的串转化为4进制数，然后找到重复的数字。

代码：

```cpp
#include <set>
class Solution {
public:
    vector<string> findRepeatedDnaSequences(string s) {
        vector<string> result;
        if(s.size() < 10)
            return result;
        int a[26];
        a['A'-'A'] = 0; a['C'-'A'] = 1; a['G'-'A'] = 2; a['T'-'A'] = 3;
        multiset<int> r;
        for(int i = 0; i < s.size() - 9; i++){
            int res = 0;
            for(int j = 0; j < 10; j++){
                res = res * 4 + a[s.at(i+j) - 'A'];
            }
            if(r.count(res) == 1)
                result.push_back(s.substr(i, 10));
            r.insert(res);
        }
        return result;
    }
};
```