PAT 乙级题目讲解：1007《素数对猜想》

题目简介

本题考察的是数论中的一个经典问题：素数对猜想。

题目定义了所谓“素数对”是指一对相邻的素数，差值为 2，例如 (3, 5)、(5, 7)、(11, 13) 等。我们需要找出所有不超过给定正整数 N 的这样的素数对个数。

数据范围：$1 \leq N < 10^5$。

样例分析

输入：

```
20
```

分析过程：

- 不超过 20 的所有素数为：2, 3, 5, 7, 11, 13, 17, 19
- 素数对为：
  - (3, 5)
  - (5, 7)
  - (11, 13)
  - (17, 19)

共有 4 对。

因此输出为：

```
4
```

解题思路

变量说明

变量名	含义
`n`	输入的正整数 N
`a[i]`	布尔数组，标记 `i` 是否为合数（`false` 表示质数）
`ans`	满足条件的素数对个数

本题求解可分为以下步骤：

Step 1：用筛法筛选出所有不超过 n 的素数

筛法思路：

- 初始时将所有数标记为“可能是质数”（即 `a[i] = false`）。
- 从 2 到 $\sqrt{n}$ 枚举每一个数 i，若 i 仍为质数（即 `a[i] == false`），则将所有 i 的倍数标记为合数（即 `a[j] = true`）。

```
for(int i = 2; i * i <= n; i++){
    if(!a[i]){
        for(int j = 2 * i; j <= n; j += i){
            a[j] = 1; // 标记为合数
        }
    }
}
```

☑ Step 2：枚举所有奇数，判断 (i, i+2) 是否为素数对

- 由于除了 2 以外所有质数都是奇数，因此我们从 3 开始，枚举每个奇数 i。
- 若 i 和 $i+2$ 都是质数（即 `a[i] == false` 且 `a[i+2] == false`），则计数器加 1。

```
for(int i = 3; i + 2 <= n; i += 2){
    if(!a[i] && !a[i + 2]){
        ans++;
    }
}
```

☑ 完整代码

```
#include <bits/stdc++.h>
using namespace std;

int n;
bool a[100005]; // a[i] = false 表示 i 是质数
int main(){
    scanf("%d", &n);
    for(int i = 2; i * i <= n; i++){
        if(!a[i]){
            for(int j = 2 * i; j <= n; j += i){
                a[j] = 1;
            }
        }
    }
    int ans = 0;
    for(int i = 3; i + 2 <= n; i += 2){
        if(!a[i] && !a[i + 2]){
            ans++;
        }
    }
    printf("%d", ans);
    return 0;
}
```

🚧 常见错误提醒

错误类型	具体表现
判断质数方法错误	使用 $2 \sim n-1$ 暴力判断质数，会超时
忘记从 3 开始枚举	由于 (2, 4) 不是素数对，应从 3 开始枚举
未判断 i+2 是否超过 n	枚举上界应为 `i + 2 <= n`
素数标记逻辑相反	错误地认为 `true` 是素数，导致答案反了

☑ 总结归纳

- 本题是典型的素数筛法应用题。
- 筛法的核心是：利用质因数性质反向标记合数。
- 巧妙地从 3 开始，步长为 2 枚举可能的素数对，提高效率。
- 注意标记逻辑中 `false` 才代表质数，这是很多初学者易错之处。

🧠 思维拓展

- 素数对猜想（孪生素数猜想）：猜想存在无限多对差为 2 的素数对，至今尚未被证明。
- 本题是该猜想的有限模型：在范围 $[1, n]$ 中寻找所有“孪生素数”。
- 可进一步探究欧拉筛优化、线性筛等更高效的素数判定算法。