EDU

CSDN学院　IT实战派

比特币和区块链的相关技术细节

哈希函数

讲师：康烁

本节目标

1. 了解哈希函数的原理
2. 理解哈希函数的基本性质

什么是哈希函数?

哈希函数(Hash function)是具有下面特点的一组函数f(x):

可接受任意长度字符串作为输入

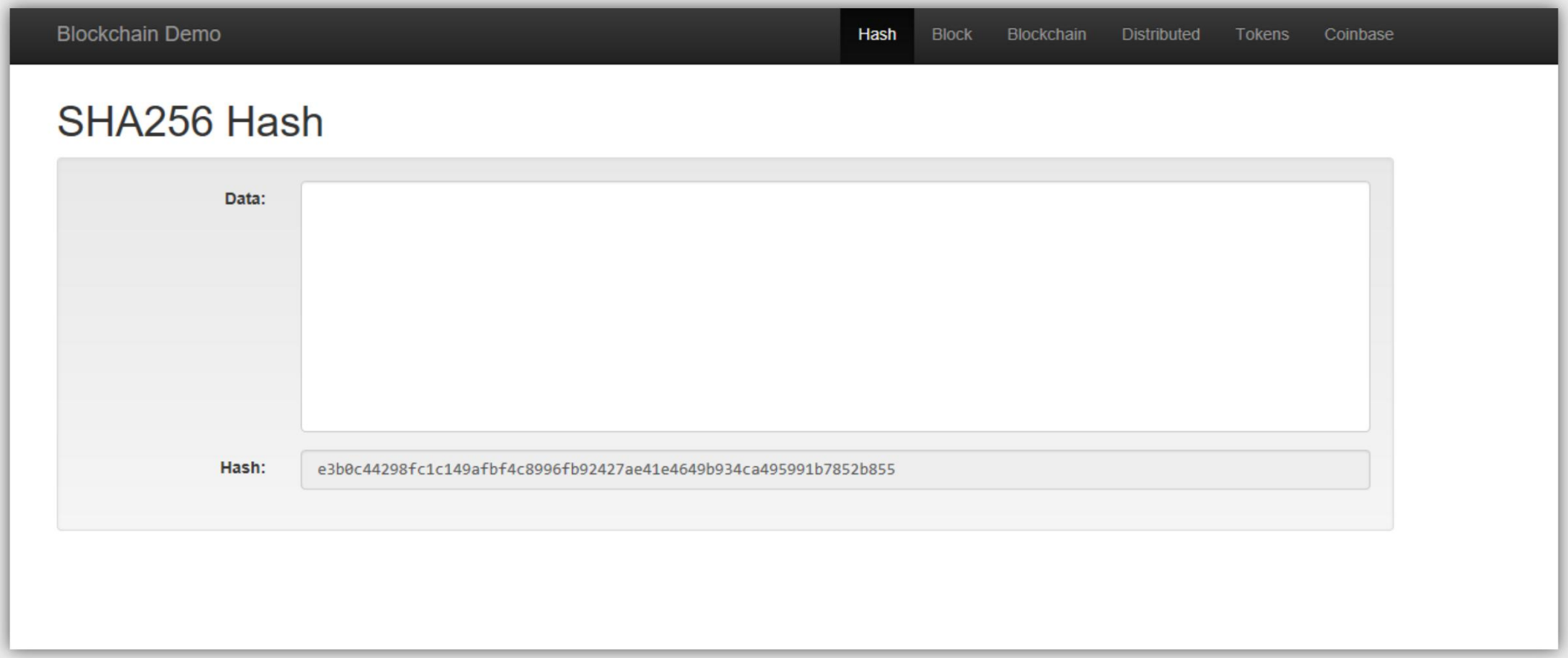
输出固定长度字符串 (在比特币中为256位，32字节)

可高效计算

不可逆运算

DEMO：哈希函数

下载地址：https://anders.com/blockchain/hash.html



哈希函数的三个性质

哈希安全性质:

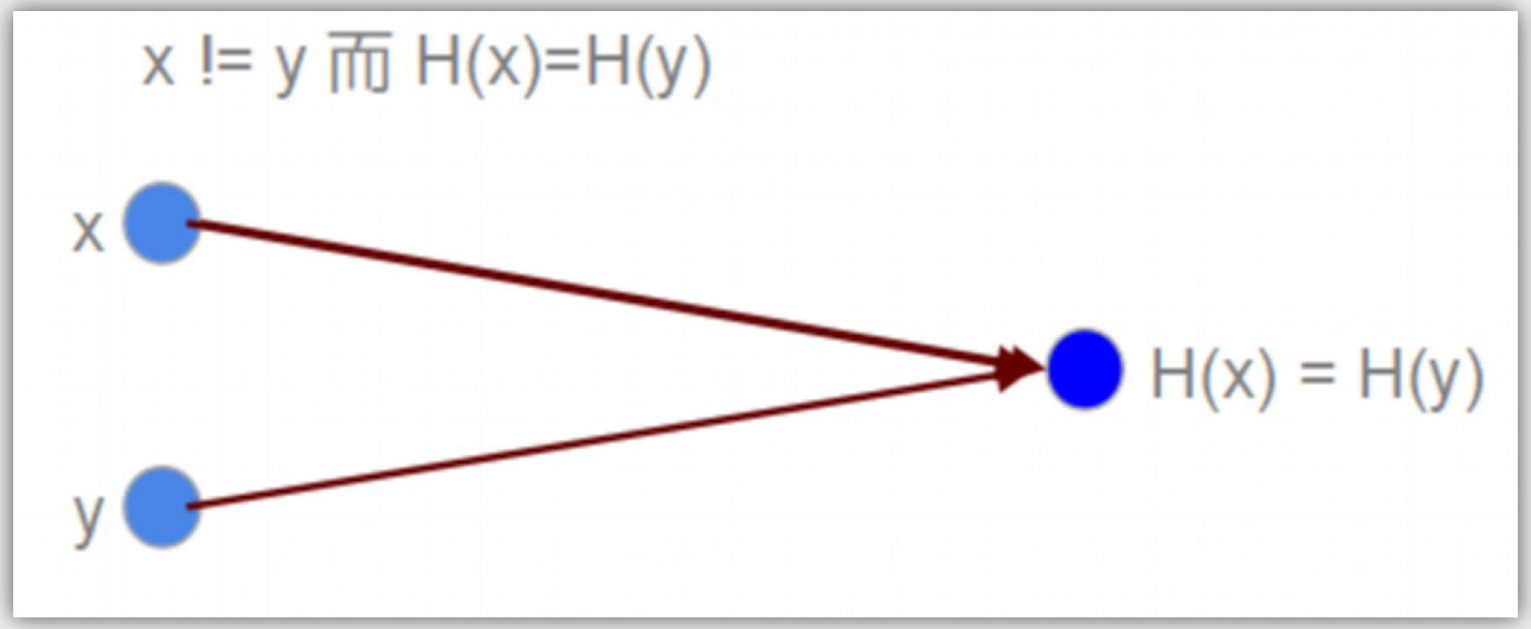
collision-free 防碰撞

hiding 信息隐藏

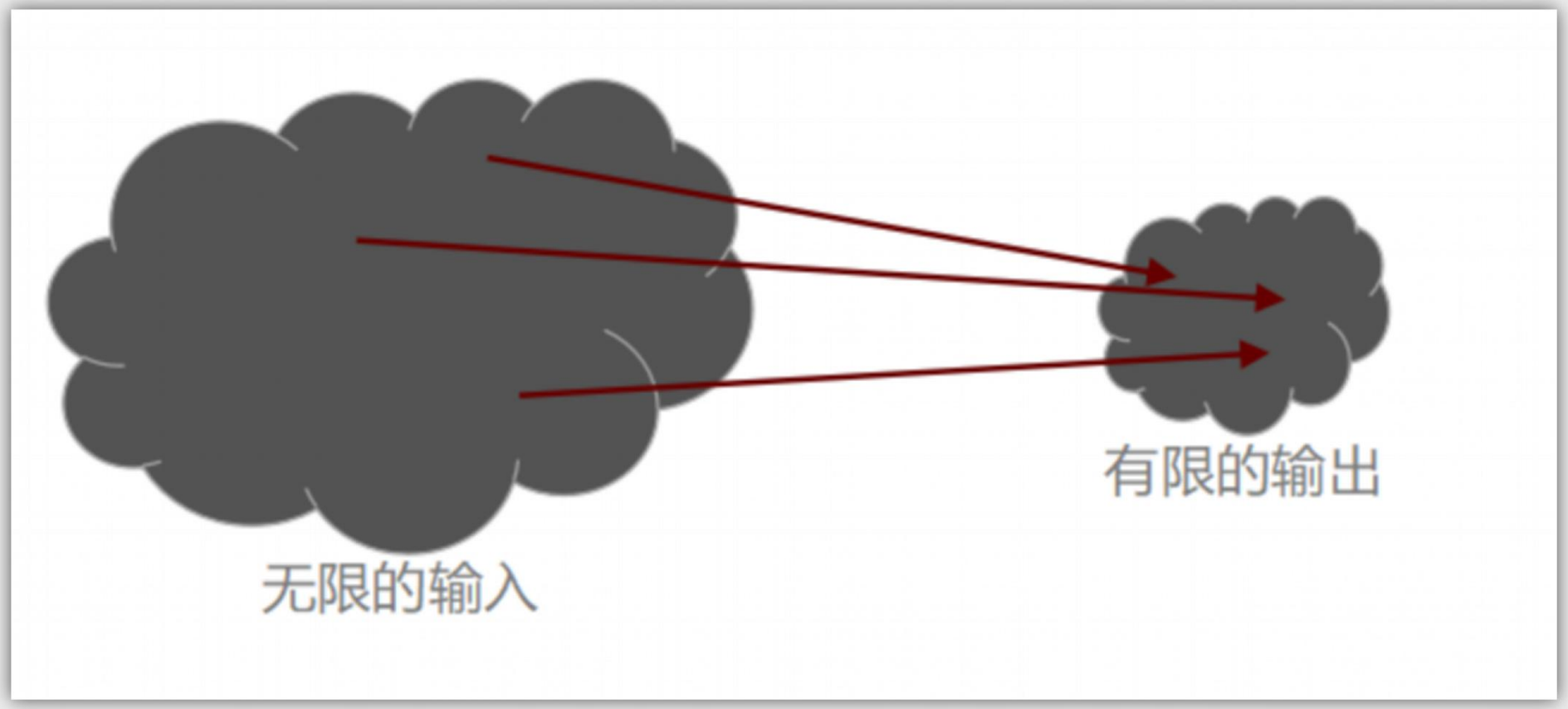
puzzle-friendly 易出难题

哈希函数性质1：防碰撞

哈希碰撞



碰撞肯定存在，但几乎找不到



问题是谁能找到

找到碰撞的方法

1. 蛮力碰撞算法

对于目前区块链中使用的哈希函数来说，还没有找到比蛮力方法更快的碰撞检测算法

2. 因此，事实上哈希算法是防碰撞的（不可逆运算；输出空间巨大）

防碰撞的应用：用哈希函数进行信息摘要

1. 因为哈希函数防碰撞，所以
2. 如果 H(x) = H(y)，则几乎可以判定 x = y
3. 因此，可以将一个文件的哈希值作为其指纹
4. 两个文件的哈希值相同，其内容必定相同，恰如两枚指纹相同，则必定来自同一人
5. 哈希值长度固定，一般为 160 bit – 512 bit，相比于文件体积来说很小

哈希函数性质 2: 信息隐藏

1. 信息隐藏的基本要求——不可逆推
2. 若 h = H(x)
3. 不可能从 h 值倒推 x

有密钥的哈希函数信息隐藏性质

1. 想隐藏信息x，先随机选取 r，得到 h = H(r | x)，无法从 h 中推测出 x（式中 | 表示字符串链接）
2. 这种信息隐藏对于 r 的随机性要求非常高，必须采用密码学意义上的真随机数，如

```go
// Go
c := 10
r := make([]byte, c)
_, err := rand.Read(b)
if err != nil {
    fmt.Println("error:", err)
    return
}
```

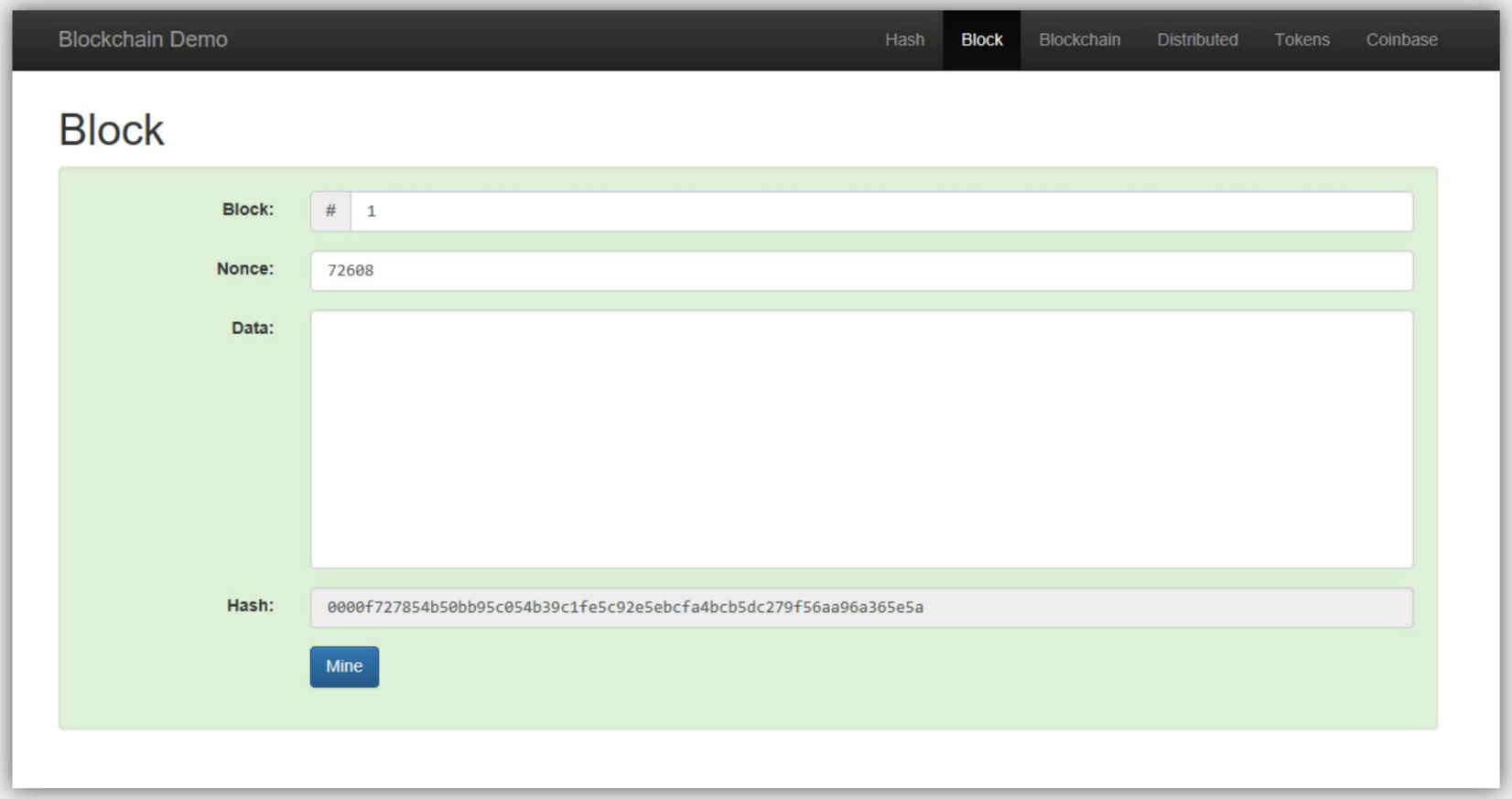
```python
# Python 3.6+
import secrets
r = secrets.token_bytes(10)

# Python 2.7
import os
r = os.urandom(10)
```

哈希函数性质 3: 易出难题

1. 随机选取 k，对于任意一个输出值 y 来说：很难找到一个值 x，使得 H(k | x) = y
2. 而想找到这样的x，没有比随机瞎猜 x 更快的算法

DEMO：用哈希函数出难题

演示地址：https://anders.com/blockchain/block.html



请找到一个 Nonce 值，使得在给定 Data 输入之下，有:

H(Nonce | Data |Block) 以 n 个 0 开头 （即小于某个阈值）

总结

1. 哈希函数的定义
2. 哈希函数的三个性质

布置作业

- 必做内容：
- 理解哈希函数的基本性质

EDU

CSDN学院　IT实战派