

# Crypto 4: Public Key



# Administrivia!

- Project 1 due Friday at 11:59 PM Pacific
  - Reminder, you have slip days if you need them
- Homework 2 due Monday at 11:59 PM Pacific
  - Reminder, slip days do not apply
- Reminder:
  - Zoom chat for conversation
  - Zoom Q&A for Questions & Answers

# Twitter Fight: Nick Vs Rust Rand\_Core Random Number Generators

- Rust (well, the 3rd party library for it) has an interface for "secure" Random Number Generators... But they aren't actually secure!
- EG, "ChaCha8Rng"
  - A **reduced round** stream cipher!
  - That has no update() function: no way of adding in entropy after seeding
  - And seed() takes only 32B total (no combining entropy!)
  - Oh, and no rollback resistance either
- **NONE** of the "Secure" RNGs are actually cryptographically secure...
  - Because none accept and consume arbitrarily long seeds or have an update to mix in more entropy
- When I say ONLY use HMAC\_DRBG, I mean it!
  - Use /dev/urandom and everything else you can think of to shove into HMAC\_DRBG

# And Vuln of the Day: CVE-2019-16303

- If you wrote an app in JHipster last year or before...
  - You probably want a password reset function...
- Password reset generates "random" URLs
  - But of course, they used a bad RNG!
- So generate a password request for your account
  - You get the RNGs state in the reset URL
- Now you can generate more password resets...
  - And predict what the "random" URL is...  
and take over any account you want!

# Public Key...

- All our previous primitives required a "miracle":
  - We somehow have to have Alice and Bob get a shared  $k$ .
- Enter Public Key cryptography: the miracle of modern cryptography
  - How starting Friday, but *what* today
- Three primitives:
  - Public Key Agreement
  - Public Key Encryption
  - Public Key Signatures
- Based on some families of magic math...
  - For us, we will use some group-theory based primitives



# Public Key Agreement

- Alice and Bob have a channel...
  - There may be an eavesdropper *but not a manipulator*
- The goal: Alice & Bob agree on a *random* value
  - This will be *k* for all subsequent communication
- When done, the key is thrown away
  - Designed to prevent an attacker who later recovers Alice or Bob's long lived secrets from finding *k*.

# Public Key Encryption

- Alice has **two** keys:
  - $K_{pub}$ : Her public key, anyone can know
  - $K_{priv}$ : Her private key, a deep dark secret
- Anyone has access to Alice's public key
- For anyone to send a message to Alice:
  - Create a random session key  $k$ 
    - Used to encrypt the rest of the message
  - Encrypt  $k$  using Alice's  $K_{pub}$ .
- Only Alice can **decrypt** the message
  - The decryption function only works with  $K_{priv}$ !

# Public Key Signatures

- Once again, Alice has **two** keys:
  - $K_{pub}$ : Her public key, anyone can know
  - $K_{priv}$ : Her private key, a deep dark secret
- She can sign a message
  - Calculate  $H(M)$
  - $S(K_{priv}, H(M))$ : Sign  $H(M)$  with  $K_{priv}$ .
- Anyone can now verify
  - Recalculate  $H(M)$
  - $V(K_{pub}, S(K_{priv}, H(M)), H(M))$ : Verify that the signature was created with  $K_{priv}$



# Things To Remember...

- Public key is ***slow!***
  - Orders of magnitude slower than symmetric key
- Public key is based on delicate magic math
  - Discrete log in a group is the most common
  - RSA
  - Some new "post-quantum" magic...
- Some systems in particular are easy to get wrong
  - We will get to some of the epic crypto-fails later

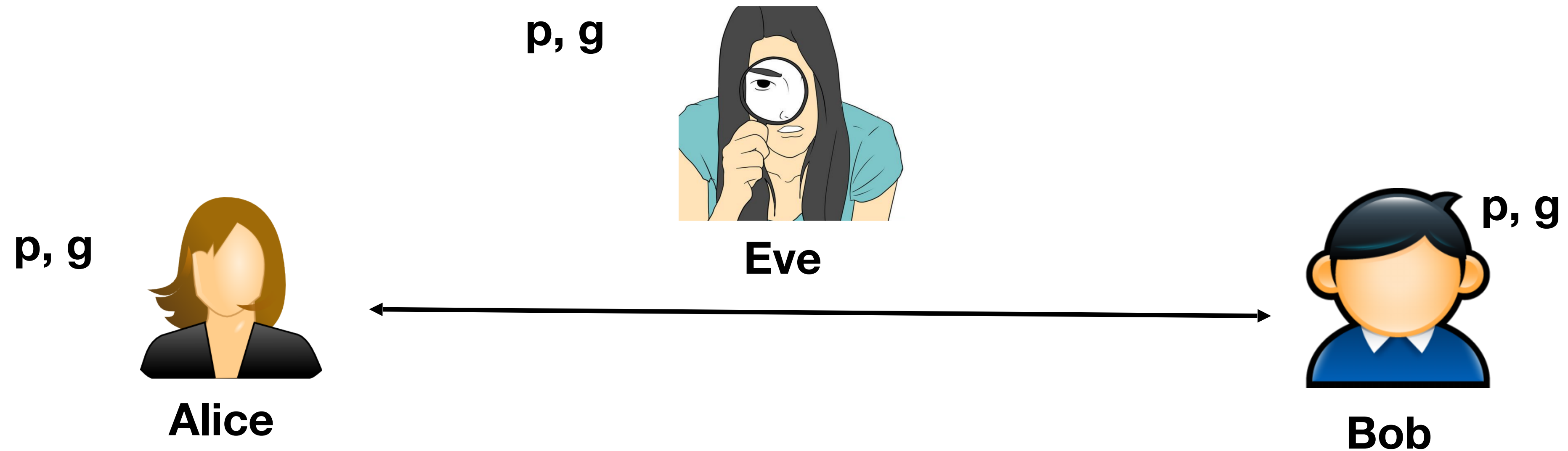
# Our Roadmap For Public Key...

- Public Key:
  - Something **everyone** can know
- Private Key:
  - The secret belonging to a specific person
- Diffie/Hellman:
  - Provides key exchange with no pre-shared secret
- ElGamal & RSA:
  - Provide a message to a recipient only knowing the recipient's **public key**
- DSA & RSA signatures:
  - Provide a message that anyone can prove was generated with a **private key**

# Diffie-Hellman Key Exchange

- What if instead they can somehow generate a random key when needed?
- Seems impossible in the presence of Eve observing all of their communication ...
  - How can they exchange a key without her learning it?
- But: actually is possible using public-key technology
  - Requires that Alice & Bob know that their messages will reach one another without any meddling
- Protocol: Diffie-Hellman Key Exchange (DHE)
  - The E is "Ephemeral", we use this to create a temporary key for other uses and then forget about it

# Diffie-Hellman Key Exchange



1. Everyone agrees in advance on a well-known (large) prime  $p$  and a corresponding  $g$ :  $1 < g < p-1$

# DHE

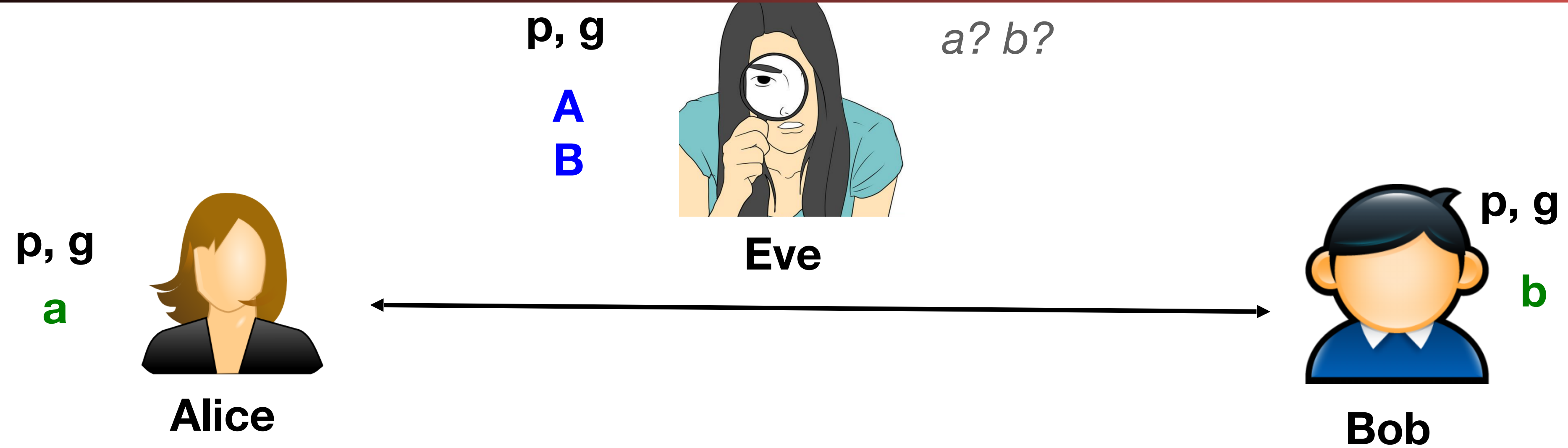


2. Alice picks **random** secret ' $a$ ':  $1 < a < p-1$

3. Bob picks **random** secret ' $b$ ':  $1 < b < p-1$

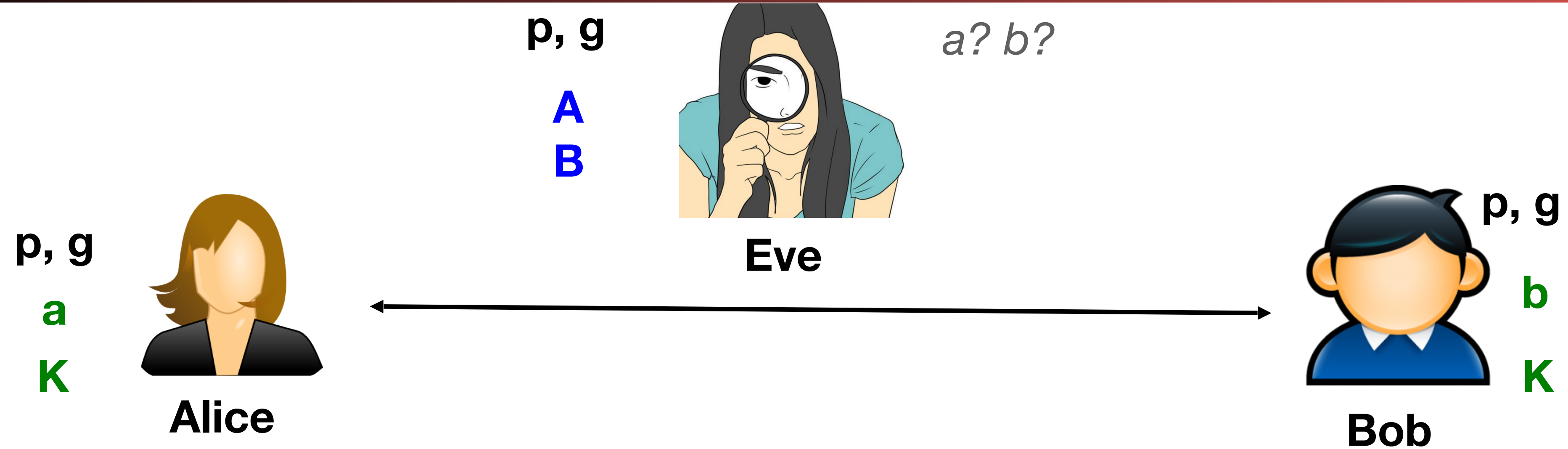


# DHE



4. Alice sends  $A = g^a \bmod p$  to Bob
5. Bob sends  $B = g^b \bmod p$  to Alice

# DHE



$$A = g^a \bmod p$$
$$B$$

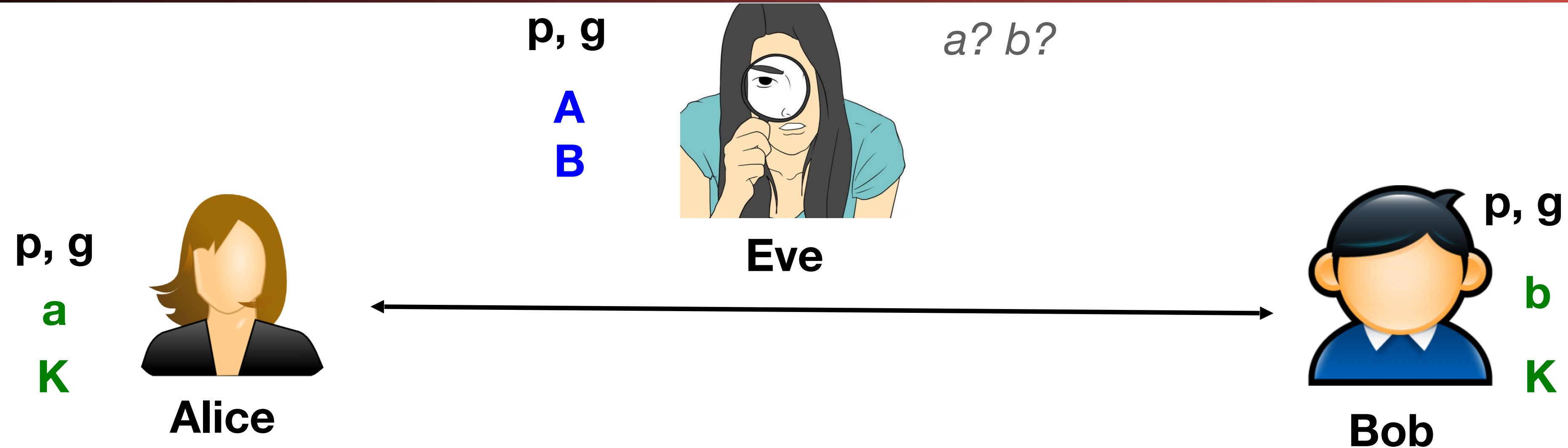
6. Alice knows  $\{a, A, B\}$ , computes  $K = B^a \bmod p = (g^b)^a = g^{ba} \bmod p$

7. Bob knows  $\{b, A, B\}$ , computes  $K = A^b \bmod p = (g^a)^b = g^{ab} \bmod p$

8.  $K$  is now the shared secret key.

$$g^b \bmod p = B$$
$$A$$

# DHE



While Eve knows  $\{p, g, g^a \bmod p, g^b \bmod p\}$ , believed to be **computationally infeasible** for her to then deduce  $K = g^{ab} \bmod p$ .

She can easily construct  $A \cdot B = g^a \cdot g^b \bmod p = g^{a+b} \bmod p$ .

But computing  $g^{ab}$  requires ability to take *discrete logarithms* mod  $p$ .  
Discrete log over the group defined by  $p$  and  $g$  **presumed** to be hard

# This is Ephemeral Diffie/Hellman

- **$K = g^{ab} \bmod p$**  is used as the basis for a "session key"
  - A symmetric key used to protect subsequent communication between Alice and Bob
    - In general, public key operations are vastly more expensive than symmetric key, so it is mostly used just to agree on secret keys, transmit secret keys, or sign hashes
  - If either **a** or **b** is random, **K** is random
- When Alice and Bob are done, they discard **K**, **a**, **b**
  - This provides **forward secrecy**: Alice and Bob don't retain any information that a later attacker who can compromise Alice or Bob's secrets could use to decrypt the messages exchanged with **K**.
  - This is also why it is called "Ephemeral" D/H

# Diffie Hellman is part of more generic problem

- This involved deep mathematical voodoo called "Group Theory"
  - Its actually done under a group  $G$
- Two main groups of note:
  - Numbers mod  $p$  with generator  $g$
  - Point addition in an elliptic curve  $C$ 
    - Usually identified by number, eg. p256, p384 (NSA-developed curves) or Curve25519 (developed by Dan Bernstein, also 256b long)
- So EC (Elliptic Curve) == different group
  - Thought to be harder so fewer bits: 384b ECDHE ?= 3096b DHE
  - But still not as hard as AES: 128b AES ?= 256b ECDHE ?= 2048b DHE
  - But otherwise, its "add EC to the name" for something built on discrete log

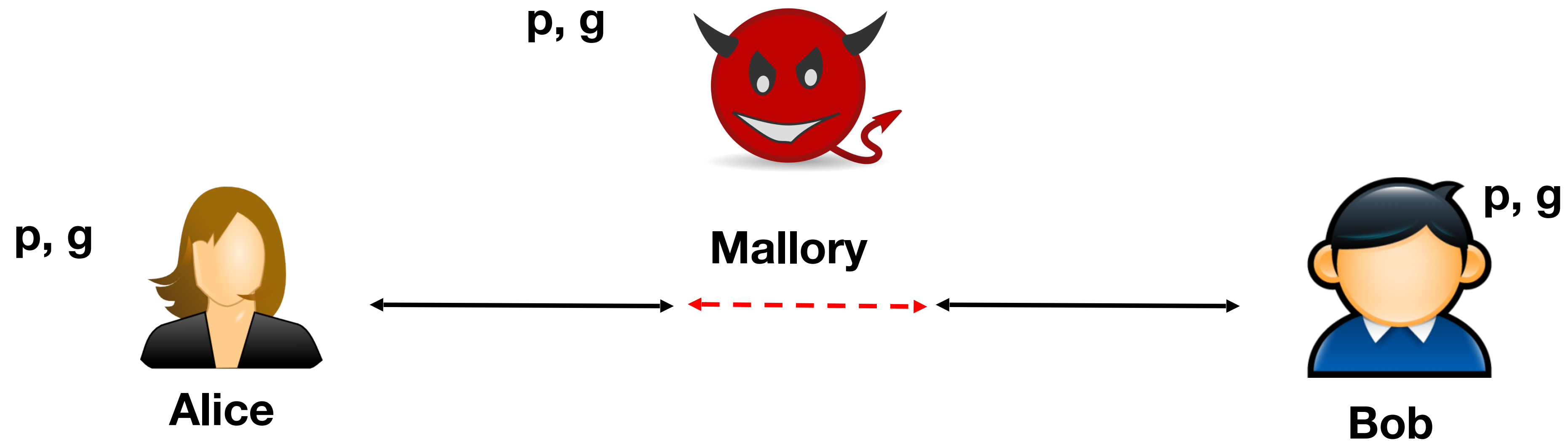


# But Its Not That Simple

- What if Alice and Bob aren't facing a passive eavesdropper
  - But instead are facing Mallory, an **active** Man-in-the-Middle
- Mallory has the ability to change messages:
  - Can remove messages and add his own
- Lets see... Do you think DHE will still work as-is?

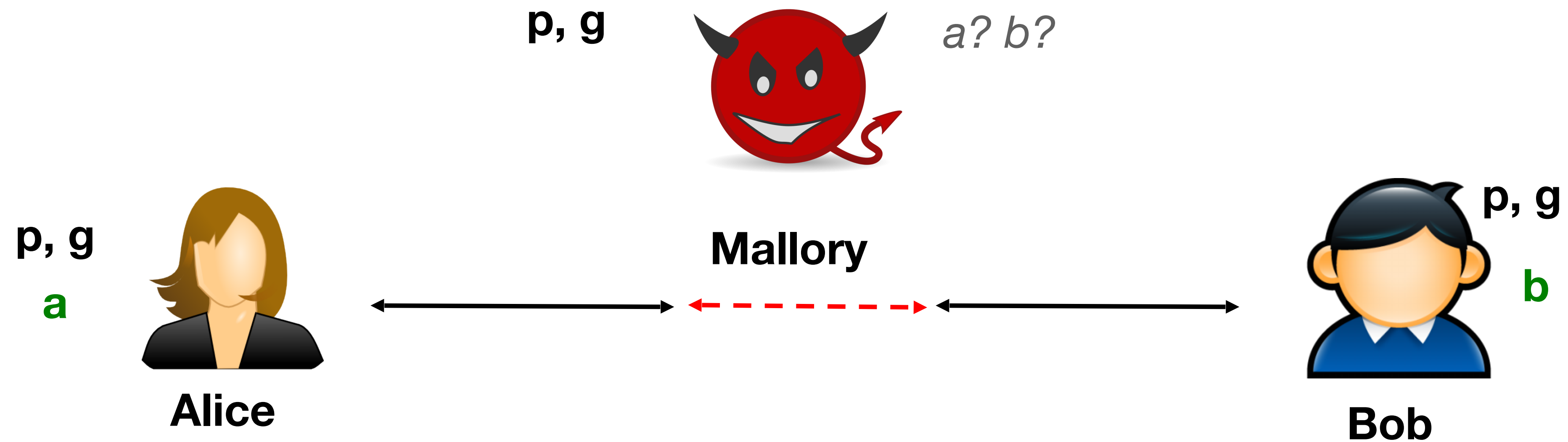


# Attacking DHE as a MitM



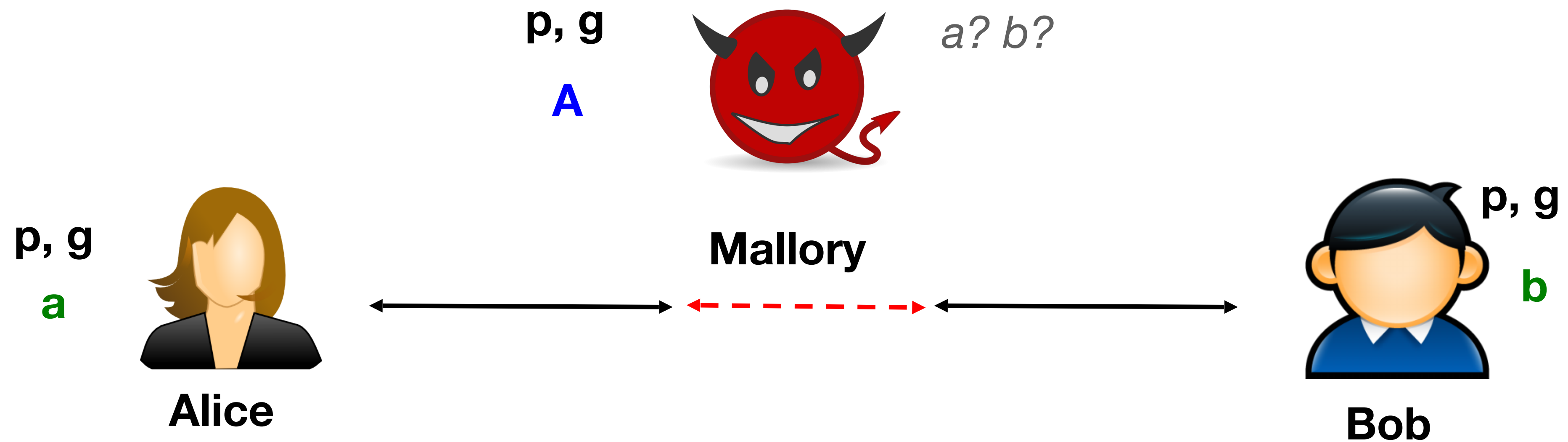
What happens if instead of Eve watching, Alice & Bob face the threat of a hidden Mallory (MITM)?

# The MitM Key Exchange



2. Alice picks **random** secret ' $a$ ':  $1 < a < p-1$

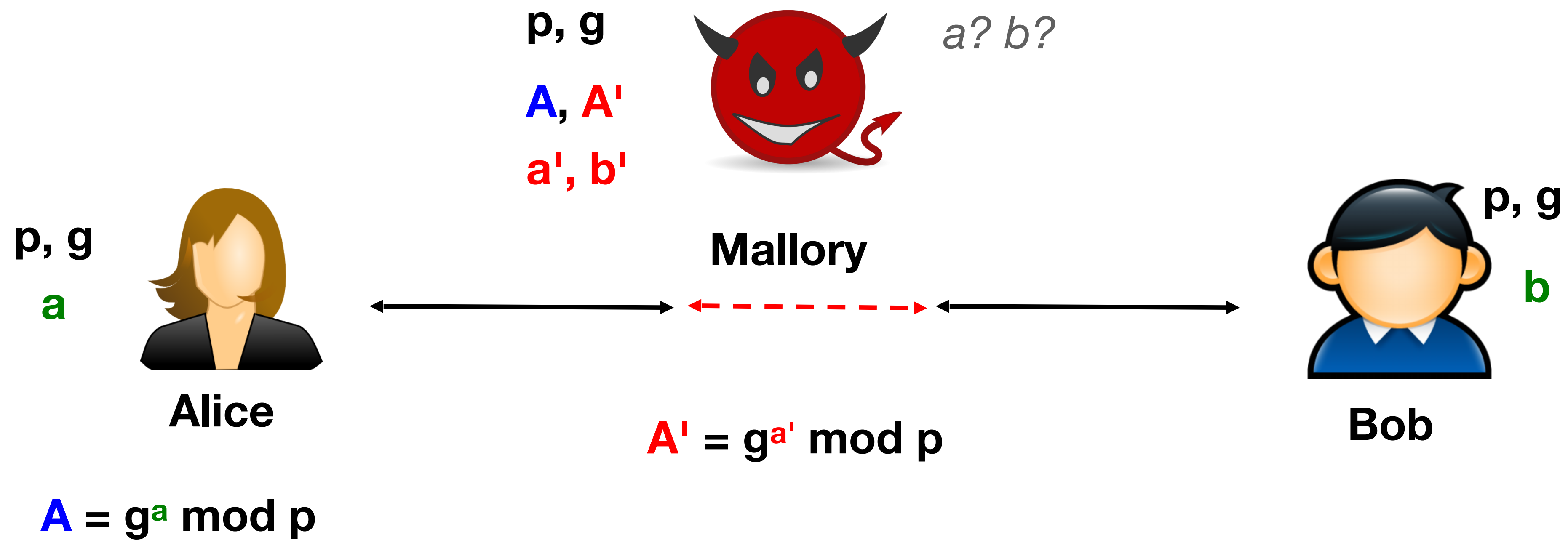
3. Bob picks **random** secret ' $b$ ':  $1 < b < p-1$



$$A = g^a \bmod p$$

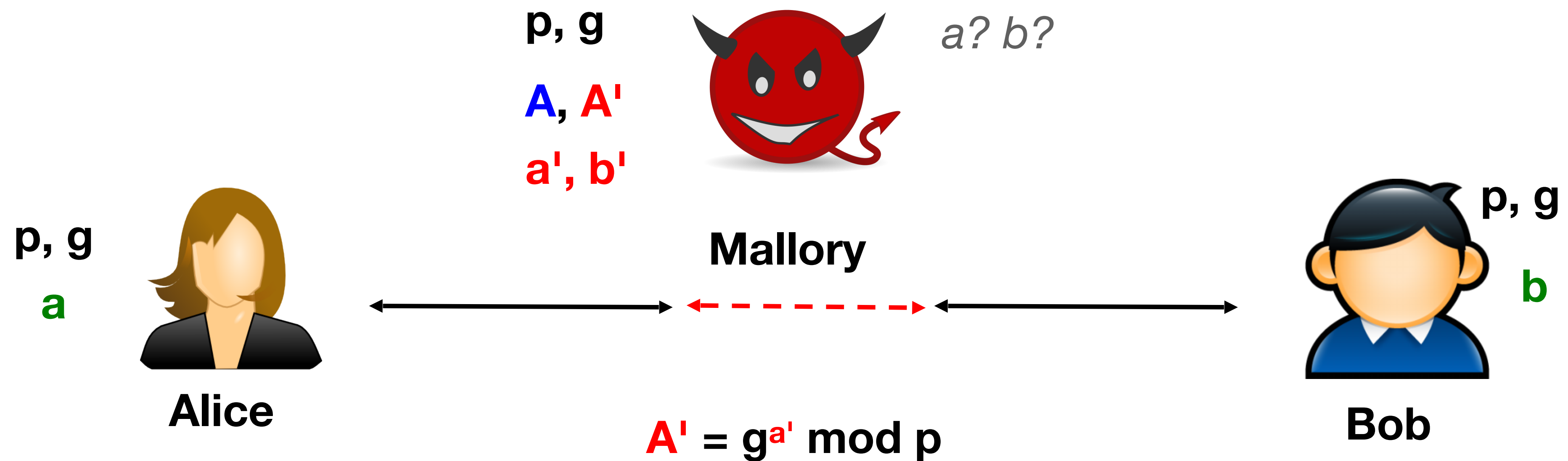
4. Alice sends  $A = g^a \bmod p$  to Bob

5. Mallory prevents Bob from receiving  $A$



6. Mallory generates her own  $a', b'$
7. Mallory sends  $A' = g^{a'} \bmod p$  to Bob

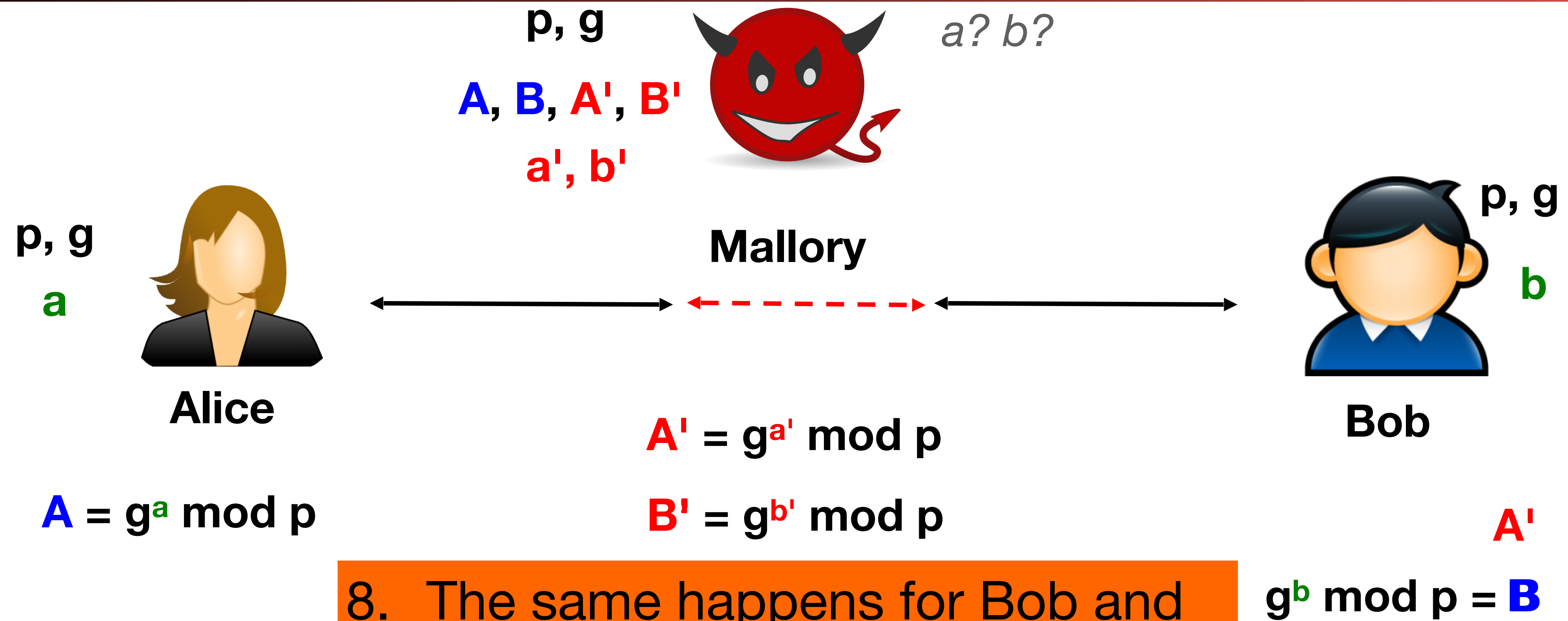




$$A = g^a \text{ mod } p$$

8. The same happens for Bob and  $B/B'$

$$g^b \text{ mod } p = B$$



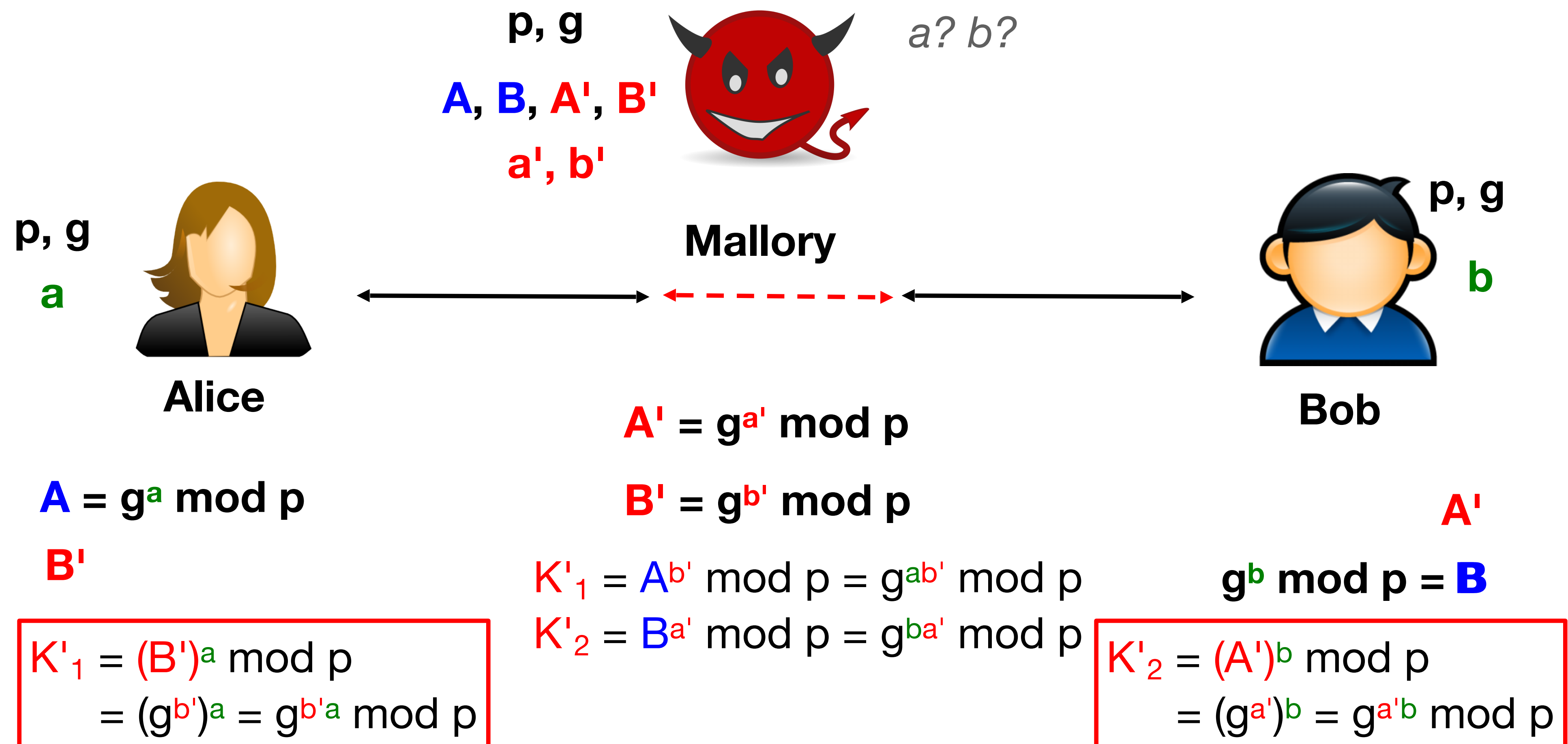
8. The same happens for Bob and  $B/B'$

$$g^b \bmod p = B$$

9. Alice and Bob now compute keys they share with ...  
Mallory!

10. Mallory can relay encrypted traffic between the two ...

10'. Modifying it or making stuff up *however she wishes*



# So We Will Want More...

- This is online:
  - Alice and Bob actually need to be active for this to work...
- So we want offline encryption:
  - Bob can send a message to Alice that Alice can read at a later date
- And authentication:
  - Alice can publish a message that Bob can verify was created by Alice later
  - Can also be used as a building-block for eliminating the MitM in the DHE key exchange:  
Alice authenticates **A**, Bob verifies that he receives **A** not **A'**.

# Public Key Cryptography #1: RSA

- Alice generates two *large* primes, **p** and **q**
  - They should be generated randomly:  
Generate a large random number and then use a "primality test":  
A *probabilistic* algorithm that checks if the number is prime
- Alice then computes **n = p\*q** and  **$\phi(n) = (p-1)(q-1)$** 
  - **$\phi(n)$**  is Euler's totient function, in this case for a composite of two primes
- Chose random  **$2 < e < \phi(n)$** 
  - **e** also needs to be relatively prime to  **$\phi(n)$**  but it can be small
- Solve for  **$d = e^{-1} \bmod \phi(n)$** 
  - You can't solve for **d** without knowing  **$\phi(n)$** , which requires knowing **p** and **q**
- **n, e** are public, **d, p, q, and  $\phi(n)$**  are secret



# RSA Encryption

- Bob can easily send a message  $m$  to Alice:
  - Bob computes  $c = m^e \bmod n$
  - Without knowing  $d$ , it is believed to be intractable to compute  $m$  given  $c$ ,  $e$ , and  $n$ 
    - But if you can get  $p$  and  $q$ , you can get  $d$ :  
It is ***not known*** if there is a way to compute  $d$  without also being able to factor  $n$ , but it is known that if you can factor  $n$ , you can get  $d$ .
    - And factoring is ***believed*** to be hard to do
- Alice computes  $m = c^d \bmod n = m^{ed} \bmod n$
- Time for some math magic...

# RSA Encryption/Decryption, con't

- So we have:  $D(C, K_D) = (M^{e \cdot d}) \bmod n$
- Now recall that  $d$  is the **multiplicative inverse** of  $e$ , modulo  $\phi(n)$ , and thus:  
 $e \cdot d = 1 \bmod \phi(n)$  (by definition)  
 $e \cdot d - 1 = k \cdot \phi(n)$  for some  $k$
- Therefore  $D(C, K_D) = M^{e \cdot d} \bmod n = (M^{e \cdot d - 1}) \cdot M \bmod n$   
 $= (M^{k \phi(n)}) \cdot M \bmod n$   
 $= [(M^{\phi(n)})^k] \cdot M \bmod n$   
 $= (1^k) \cdot M \bmod n$  *by Euler's Theorem:  $a^{\phi(n)} \bmod n = 1$*   
 $= M \bmod n = M$

(believed) Eve can recover  $M$  from  $C$  *iff* Eve can factor  $n=p \cdot q$

# But It Is Not That Simple...

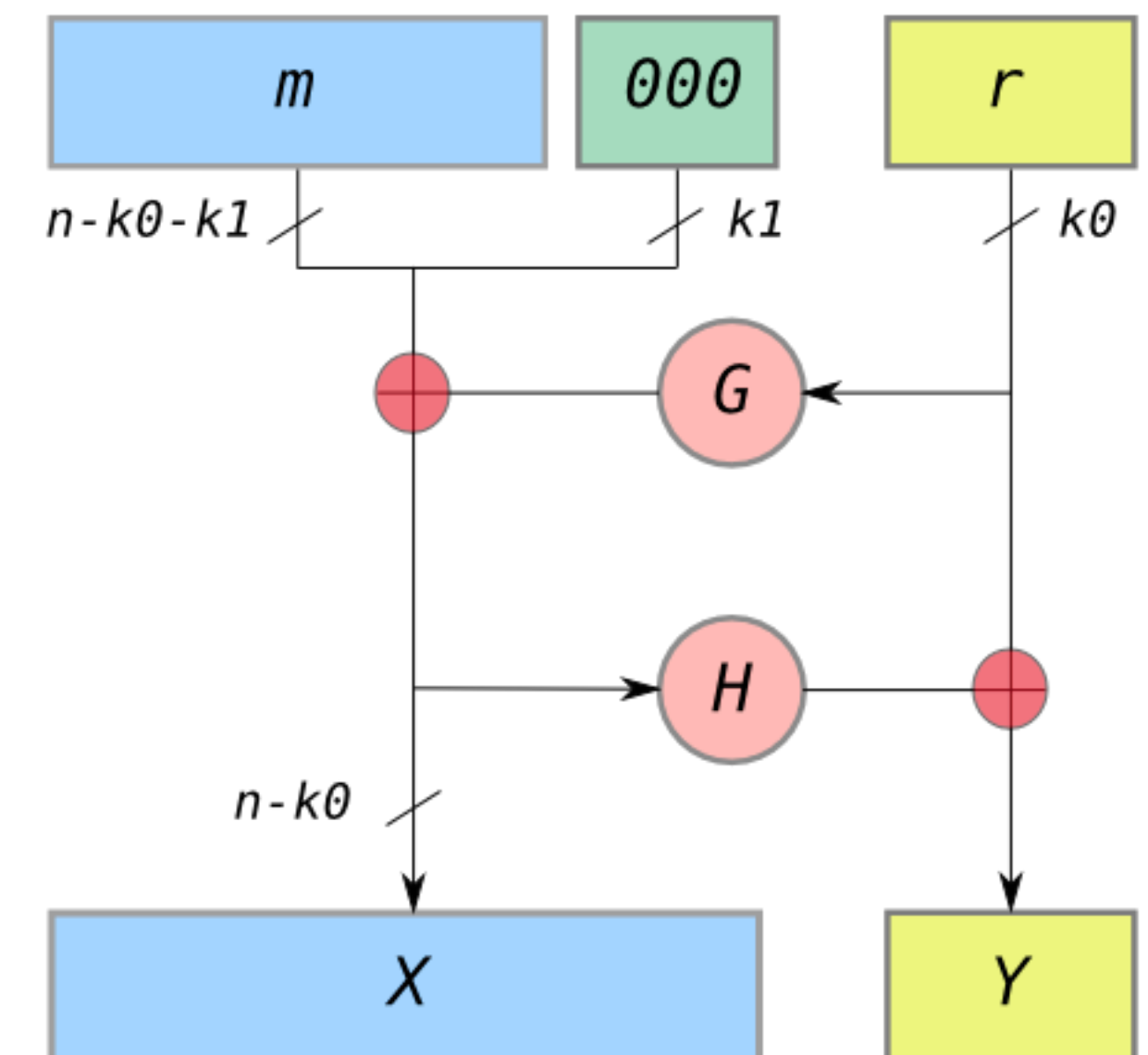
- What if Bob wants to send the same message to Alice twice?
  - Sends  $m^{e_a} \bmod n_a$  and then  $m^{e_a} \bmod n_a$
  - Oops, not IND-CPA!
- What if Bob wants to send a message to Alice, Carol, and Dave:
  - $m^{e_a} \bmod n_a$   
 $m^{e_b} \bmod n_b$   
 $m^{e_c} \bmod n_c$
  - This ends up leaking information an eavesdropper can use *especially* if  $3 = e_a = e_b = e_c$ !
- Oh, and problems if both  $e$  and  $m$  are small...
- As a result, you ***can not*** just use plain RSA:
  - You need to use a "padding" scheme that makes the input random but reversible



# RSA-OAEP

## (Optimal asymmetric encryption padding)

- A way of processing  $m$  with a hash function & random bits
  - Effectively "encrypts"  $m$  replacing it with  $X = [m, 0...] \oplus G(r)$ 
    - $G$  and  $H$  are hash functions (EG SHA-256)  
 $k_0 = \#$  of bits of randomness,  $\text{len}(m) + k_1 + k_0 = n$
  - Then replaces  $r$  with  $Y = H(G(r) \oplus [m, 0...]) \oplus R$
  - This structure is called a "Feistel network":
    - It is always designed to be reversible.  
Many block ciphers are based on this concept applied multiple times with  $G$  and  $H$  being functions of  $k$  rather than just fixed operations
- This is more than just block-cipher padding (which involves just adding simple patterns)
  - Instead it serves to both pad the bits and make the data to be encrypted "random"





# But Its Not That Simple...

## Timing Attacks

- Using normal math, the **time** it takes for Alice to decrypt **c** depends on **c** and **d**
  - Ruh roh, this can leak information...
  - More complex RSA implementations take advantage of knowing **p** and **q** directly... but also leak timing
- People have used this to guess and then check the bits of **q** on OpenSSL
  - <http://crypto.stanford.edu/~dabo/papers/ssl-timing.pdf>
- And even more subtle things are possible...

```
x = C
for j = 1 to n
    x = mod(x2, N)
    if dj == 1 then
        x = mod(xC, N)
    end if
next j
return x
```



SwiftOnSecurity  
@SwiftOnSecurity

Following





# So How to Find Bob's Key?

- Lots of stuff later, but for now...  
The Leap of Faith!
- Alice wants to talk to Bob:
  - "Hey, Bob, tell me your public key!"
- Now on all subsequent times...
  - "Hey, Bob, tell me your public key", and check to see if it is different from what Alice remembers
- Works assuming the ***first time*** Alice talks to Bob there isn't a Man-in-the-Middle
  - ssh uses this

# RSA Signatures...

- Alice computes a hash of the message  **$H(m)$** 
  - Alice then computes  **$s = (H(m))^d \bmod n$**
- Anyone can then verify
  - **$v = s^e \bmod m = ((H(m))^d)^e \bmod n = H(m)$**
- Once again, there are "F-U"s...
  - Have to use a proper encoding scheme to do this properly and all sort of other traps
  - One particular trap: a scenario where the attacker can get Alice to repeatedly sign things (an "oracle")



# But Signatures Are Super Valuable...

- They are how we can prevent a MitM!
- If Bob knows Alice's key, and Alice knows Bob's...
  - How will be "next time"
- Alice doesn't just send a message to Bob...
  - But creates a random key  $k$ ...
  - Sends  $E(M, K_{\text{sess}})$ ,  $E(K_{\text{sess}}, B_{\text{pub}})$ ,  $S(H(M), A_{\text{priv}})$
- Only Bob can decrypt the message, and Bob can verify the message came from Alice
  - So Mallory is SOL!

# RSA Isn't The Only Public Key Algorithm

- Isn't RSA enough?
  - RSA isn't particularly compact or efficient: dealing with 2000b (comfortably secure) or 3000b (NSA-paranoia) bit operations
  - Can we get away with fewer bits?
    - Well, Diffie-Hellman isn't any better...
    - But ***elliptic curve*** Diffie-Hellman is
- RSA also had some patent issues
  - So an attempt to build public key algorithms around the Diffie-Hellman problem



# El-Gamal

- Just like Diffie-Hellman...
  - Select **p** and **g**
    - These are public and can be shared:  
Note, they need to be carefully considered how to create p and g...  
Math beyond the level of this class
- Alice chooses **x** randomly as her private key
  - And publishes **h = g<sup>x</sup> mod p** as her public key
- Bob, to encrypt m to Alice...
  - Selects a *random* **y**, calculates **c<sub>1</sub> = g<sup>y</sup> mod p**, **s = h<sup>y</sup> mod p = g<sup>xy</sup> mod p**
    - **s** becomes a shared secret between Alice and Bob
  - Maps message **m** to create **m'**, calculates **c<sub>2</sub> = m' \* s mod p**
- Bob then sends {**c<sub>1</sub>**, **c<sub>2</sub>**}



# El-Gamal Decryption

- Alice first calculates  $s = c_1^x \bmod p$ 
  - Then Alice calculates  $m' = c_2 * s^{-1} \bmod p$
  - Then Alice calculates the inverse of the mapping to get  $m$
- Of course, there are problems...
  - Attacker can always change  $m'$  to  $2m'$
  - What if Bob screws up and reuses  $y$ ?
  - $c_2 = m_1' * s \bmod p$   
 $c_2' = m_2' * s \bmod p$
  - Ruh roh, this leaks information:  
 $c_2 / c_2' = m_1' / m_2'$ 
    - So if you know  $m_1...$



# In Practice: Session Keys...

- You use the public key algorithm to encrypt/agree on a session key..
  - And then encrypt the real message with the session key
  - You **never** actually encrypt the message itself with the public key algorithm
  - Often a set of keys: encrypt and MAC keys that are separate in each direction
- Why?
  - Public key is **slow**... Orders of magnitude slower than symmetric key
  - Public key may cause weird effects:
    - EG, El Gamal where an attacker can change the message to  **$2m$** ...
    - If  **$m$**  had meaning, this would be a problem
    - But if it just changes the encryption and MAC keys, the main message won't decrypt

# DSA Signatures...

- Again, based on Diffie-Hellman
  - Two initial parameters, **L** and **N**, and a hash function **H**
    - **L** == key length, eg 2048
    - **N** <= len(**H**), e.g. 256
  - An N-bit prime **q**, an L-bit prime **p** such that **p - 1** is a multiple of **q**, and **g** = **h**<sup>(p-1)/q</sup> mod **p** for some arbitrary **h** (1 < h < p - 1)
  - {**p**, **q**, **g**} are public parameters
- Alice creates her own random private key **x** < **q**
  - Public key **y** = **g**<sup>x</sup> mod **p**

# Alice's Signature...

- Create a random value  $k < q$ 
  - Calculate  $r = (g^k \bmod p) \bmod q$ 
    - If  $r = 0$ , start again
  - Calculate  $s = k^{-1} (H(m) + xr) \bmod q$ 
    - If  $s = 0$ , start again
  - Signature is  $\{r, s\}$  (Advantage over an El-Gamal signature variation: Smaller signatures)
- Verification
  - $w = s^{-1} \bmod q$
  - $u_1 = H(m) * w \bmod q$
  - $u_2 = r * w \bmod q$
  - $v = (g^{u_1} y^{u_2} \bmod p) \bmod q$
  - Validate that  $v = r$



# But Easy To Screw Up...

- **k** is not just a nonce... It must be random and **secret**
  - If you know **k**, you can calculate **x**
- And even if you just reuse a random **k**... for two signatures **s<sub>a</sub>** and **s<sub>b</sub>**
  - A bit of algebra proves that  $k = (H_A - H_B) / (s_a - s_b)$
- A good reference:
  - How knowing k tells you x:  
<https://rdist.root.org/2009/05/17/the-debian-gpg-disaster-that-almost-was/>
  - How two signatures tells you k:  
<https://rdist.root.org/2010/11/19/dsa-requirements-for-random-k-value/>





# And ***NOT*** theoretical: Sony Playstation 3 DRM

- The PS3 was designed to only run ***signed*** code
  - They used ECDSA as the signature algorithm
  - This prevents unauthorized code from running
  - They had an ***option*** to run alternate operating systems (Linux) that they then removed
- Of course this was catnip to reverse engineers
  - Best way to get people interested: ***remove*** Linux from a device...
- It turns for out one of the key authentication keys used to sign the firmware...
  - Ended up reusing the same k for multiple signatures!





# And **NOT** Theoretical: Android RNG Bug + Bitcoin

- OS Vulnerability in 2013  
Android "SecureRandom" wasn't actually secure!
  - Not only was it low entropy, it would occasionally return the **same value multiple times**
- Multiple Bitcoin wallet apps on Android were affected
  - "Pay B Bitcoin to Bob" is signed by Alice's public key using ECDSA
    - Message is broadcast publicly for all to see
  - So you'd have cases where "Pay B to Bob" and "Pay C to Carol" were signed with the same **k**
- So **of course** someone scanned for **all** such Bitcoin transactions





# And *Still* Happens!

## Chromebook

- Chromebooks have a built in U2F "Security key"
  - Enables signatures using 256b ECDSA to validate to particular websites
- There was a bug in the secure hardware!
  - Instead of using a random  $k$  that was 256b long, a bug caused it to be 32b long!
  - So an attacker who had a signature could simply try all possible  $k$  values!
- Fortunately in this case the damage was slight: this is for authenticating to a single website: each site used its own private key
- But still...
- <https://www.chromium.org/chromium-os/u2f-ecdsa-vulnerability>



# So What To Use?

- Paranoids like me:  
Good libraries and use the parameters from NSA's CNSA suite
  - Open algorithms approved for Top Secret communication
  - Better yet, libraries that implement full protocols that use these under the hood!
- Symmetric cipher: AES: 256b
  - CFB mode, thankyouverymuch. Counter mode and modes which include counter mode can DIAF...
- Hash function: SHA-384
  - Use HMAC for MAC
- RSA: 3072b
- Diffie/Hellman: 3072b
- ECDH/ECDSA: P-384
  - But really, this is extra paranoid, 2048b RSA/DH, 256b EC, 128b AES, SHA-256 excellent in practice

# How Can We Communicate With Someone New?

- Public-key crypto gives us amazing capabilities to achieve confidentiality, integrity & authentication without shared secrets ...
- But how do we solve MITM attacks?
- How can we trust we have the true public key for someone we want to communicate with?
- Ideas?

# Trusted Authorities

- Suppose there's a party that everyone agrees to trust to confirm each individual's public key
  - Say the Governor of California
- Issues with this approach?
  - How can everyone agree to trust them?
  - Scaling: huge amount of work; single point of failure ...
    - ... and thus Denial-of-Service concerns
  - How do you know you're talking to the right authority??





# Trust Anchors

- Suppose the trusted party distributes their key so everyone has it ...













Gavin Newsom's Public Key is  
0x6a128b3d3dc67edc74d690b19e072f64





# Trust Anchors

- Suppose the trusted party distributes their key so everyone has it ...
- We can then use this to bootstrap trust
  - As long as we have confidence in the decisions that that party makes



# Digital Certificates

- Certificate (“cert”) = signed claim about someone’s public key
  - More broadly: a signed *attestation* about some claim
- Notation:
  - $\{ M \}_K$  = “message M encrypted with public key k”
  - $\{ M \}_{K^{-1}}$  = “message M signed w/ private key for K”
- E.g. M = “Nick's public key is  $K_{\text{Nick}} = 0xF32A99B...$ ”  
Cert: M,
  - $\{ \text{“Nick's public key ... } 0xF32A99B... \text{”} \}_{K^{-1}}_{\text{Gavin}}$
  - $= 0x923AB95E12...9772F$



# *Certificate*



Gavin Newsom hearby asserts:  
Nick's public key is  $K_{Nick} = \mathbf{0xF32A99B...}$   
The signature for this statement using  
 $K_{Gavin}^{-1}$  is  $\mathbf{0x923AB95E12...9772F}$



# *Certificate*



Gavin Newsom hearby asserts:  
Nick's public key is  $K_{\text{Nick}} = \mathbf{0xF32A99B...}$   
The signature for this statement using  
 $K^{-1}$  **This** is  $\mathbf{0x923AB95E12...9772F}$



# Certificate



Gavin Newsom hereby asserts:

Nick's public key is  $K_{\text{Nick}} = 0xF32A99B...$

The signature for this is computed over all of this

$K_{\text{Gavin}}^{-1}$  is  $0x923AB95E12...9772F$



# Certificate



Gavin Newsom hereby asserts:  
Nick's public key is  $K_{\text{Nick}} = \mathbf{0xF32A99B...}$   
The signature for this statement using  
 $K_{\text{Gavin}}^{-1}$  is  $\mathbf{0x923AB95E12...9772F}$

and can be  
*validated* using:



# Certificate



This:

Gavin Newsom hereby asserts:

Nick's public key is  $K_{\text{Nick}}$

The signature for this statement

$K_{\text{Gavin}}^{-1}$  is **0x923AB95**

