Lecture Outline

1. The Power of Randomness
   - Reed-Solomon Fingerprinting
   - Freivalds' Protocol for Verifying Matrix Products
2. Definition of Interactive Proofs
3. Technical Concepts: low-degree extensions

The Power of Randomness: A Demonstration

EQUALITY Testing

Alice

Bob

$\boldsymbol{a} = (a_1, \ldots, a_n) \in \{0,1\}^n$

$\boldsymbol{b} = (b_1, \ldots, b_n) \in \{0,1\}^n$

Alice and Bob's Goal: Determine whether $\boldsymbol{a} = \boldsymbol{b}$, while exchanging as few bits as possible.

EQUALITY Testing

Alice

Bob

$\boldsymbol{a} = (a_1, \dots, a_n) \in \{0,1\}^n$

$\boldsymbol{b} = (b_1, \dots, b_n) \in \{0,1\}^n$

Trivial solution: Alice sends $\boldsymbol{a}$ to Bob, who checks whether $\boldsymbol{a} = \boldsymbol{b}$.

Communication cost is n.

EQUALITY Testing

Alice

Bob

$\boldsymbol{a} = (a_1, \ldots, a_n) \in \{0,1\}^n$

$\boldsymbol{b} = (b_1, \ldots, b_n) \in \{0,1\}^n$

Fact: Trivial solution is optimal amongst deterministic protocols.

A Logarithmic Cost Randomized Solution

Randomized EQUALITY Testing Protocol

- Notation:
  - Let $\boldsymbol{F}$ be any finite field with $|\boldsymbol{F}| \geq n^2$.
  - Interpret each a_i, b_i as elements of $\boldsymbol{F}$.
  - Let $p(x) = \sum_{i=1}^{n} a_i\, x^i$ and $q(x) = \sum_{i=1}^{n} b_i\, x^i$.

Randomized EQUALITY Testing Protocol

- Notation:
  - Let $\boldsymbol{F}$ be any finite field with $|\boldsymbol{F}| \geq n^2$.
  - Interpret each a_i, b_i as elements of $\boldsymbol{F}$.
  - Let $p(x) = \sum_{i=1}^{n} a_i\, x^i$ and $q(x) = \sum_{i=1}^{n} b_i\, x^i$.

- The Protocol:
  - Alice picks a random $r \in \boldsymbol{F}$ and sends $(r, p(r))$ to Bob.
  - Bob outputs EQUAL if $p(r) = q(r)$. Otherwise he outputs NOT-EQUAL.

Randomized EQUALITY Testing Protocol

- Notation:
  - Let $\boldsymbol{F}$ be any finite field with $|\boldsymbol{F}| \geq n^2$.
  - Interpret each a_i, b_i as elements of $\boldsymbol{F}$.
  - Let $p(x) = \sum_{i=1}^{n} a_i\, x^i$ and $q(x) = \sum_{i=1}^{n} b_i\, x^i$.

- The Protocol:
  - Alice picks a random $r \in \boldsymbol{F}$ and sends $(r, p(r))$ to Bob.
  - Bob outputs EQUAL if $p(r) = q(r)$. Otherwise he outputs NOT-EQUAL.

- Total communication: $O(\log|\boldsymbol{F}|) = O(\log n)$ bits.

Randomized EQUALITY Testing Protocol

- Notation:
  - Let $\boldsymbol{F}$ be any finite field with $|\boldsymbol{F}| \geq n^2$.
  - Interpret each a_i, b_i as elements of $\boldsymbol{F}$.
  - Let $p(x) = \sum_{i=1}^{n} a_i\, x^i$ and $q(x) = \sum_{i=1}^{n} b_i\, x^i$.

- The Protocol:
  - Alice picks a random $r \in \boldsymbol{F}$ and sends $(r, p(r))$ to Bob.
  - Bob outputs EQUAL if $p(r) = q(r)$. Otherwise he outputs NOT-EQUAL.

- Total communication: $O(\log |\boldsymbol{F}|) = O(\log n)$ bits.
- Call $p(r)$ the *Reed-Solomon fingerprint* of the vector $\boldsymbol{a}$ at r.

Correctness Analysis

- Claim 1: if $\boldsymbol{a} = \boldsymbol{b}$, then Bob outputs EQUAL with probability 1.

- Claim 2: $\boldsymbol{a} \neq \boldsymbol{b}$, then Bob outputs NOT-EQUAL with probability at least $1 - \frac{1}{n}$ over the choice of $r \in \boldsymbol{F}$.

Correctness Analysis

- Claim 1: if $\boldsymbol{a} = \boldsymbol{b}$, then Bob outputs EQUAL with probability 1.
  - Proof: Since $\boldsymbol{a} = \boldsymbol{b}$, p and q are the same polynomial, so $p(r) = q(r)$ for all $r \in \boldsymbol{F}$.

- Claim 2: $\boldsymbol{a} \neq \boldsymbol{b}$, then Bob outputs NOT-EQUAL with probability at least $1 - \frac{1}{n}$ over the choice of $r \in \boldsymbol{F}$.

Correctness Analysis

- Claim 2: $\boldsymbol{a} \neq \boldsymbol{b}$, then Bob outputs NOT-EQUAL with probability at least $1 - \frac{1}{n}$ over the choice of $r \in \boldsymbol{F}$.

Correctness Analysis

- Claim 2: $\boldsymbol{a} \neq \boldsymbol{b}$, then Bob outputs NOT-EQUAL with probability at least $1 - \frac{1}{n}$ over the choice of $r \in \boldsymbol{F}$.

FACT: Let $p \neq q$ be univariate polynomials of degree at most n. Then p and q agree on at most n inputs. Equivalently:

$$\Pr_{r \in \boldsymbol{F}}[p(r) = q(r)] \leq \frac{n}{|\boldsymbol{F}|}.$$

Correctness Analysis

- Claim 2: $\boldsymbol{a} \neq \boldsymbol{b}$, then Bob outputs NOT-EQUAL with probability at least $1 - \frac{1}{n}$ over the choice of $r \in \boldsymbol{F}$.

FACT: Let $p \neq q$ be univariate polynomials of degree at most n. Then p and q agree on at most n inputs. Equivalently:

$$\Pr_{r \in \boldsymbol{F}}[p(r) = q(r)] \leq \frac{n}{|\boldsymbol{F}|}.$$

- If $\boldsymbol{a} \neq \boldsymbol{b}$**,** then p and q are **not** the same polynomial. By **FACT**, the probability Alice picks an r such that $p(r) = q(r)$ is at most $\frac{n}{|\boldsymbol{F}|} \leq \frac{n}{n^2} \leq \frac{1}{n}$.

Main Takeaways

1. Any two distinct low-degree polynomials differ almost everywhere: if $p \neq q$ then $\Pr_{r \in \boldsymbol{F}}[p(r) = q(r)] \leq \frac{n}{|\boldsymbol{F}|}$ where n bounds the degree of p and q.
   - Corollary: If two low-degree polynomials agree at a randomly chosen input, it is "safe" to believe they are the **same** polynomial.

2. Interpreting inputs as low-degree polynomials is powerful.
   - If two inputs differ **at all**, then once interpreted as polynomials, they differ **almost everywhere**.

Freivalds' Protocol for Verifying Matrix Products

Demonstrating the Power of Randomness in Verifiable Computing

Verifying Matrix Multiplication

- Input is two matrices $\mathrm{A}, \mathrm{B} \in \boldsymbol{F}^{n \times n}$. Goal is to compute $\mathrm{A} \cdot \mathrm{B}$.
- Fastest known algorithm runs in time about $n^{2.37}$.

Verifying Matrix Multiplication

- Input is two matrices $\mathrm{A}, \mathrm{B} \in \boldsymbol{F}^{n \times n}$. Goal is to compute $\mathrm{A} \cdot \mathrm{B}$.
- Fastest known algorithm runs in time about $n^{2.37}$.
- What if an untrusted prover P claims that the answer is a matrix C? Can V **verify** that $C = \mathrm{A} \cdot \mathrm{B}$ in $O(n^2)$ time?

Verifying Matrix Multiplication

- Input is two matrices $\mathrm{A}, \mathrm{B} \in \boldsymbol{F}^{n\times n}$. Goal is to compute $\mathrm{A}\cdot\mathrm{B}$.
- Fastest known algorithm runs in time about $n^{2.37}$.
- What if an untrusted prover P claims that the answer is a matrix C? Can V **verify** that $C = \mathrm{A}\cdot\mathrm{B}$ in $O(n^2)$ time?
- Yes!

Verifying Matrix Multiplication

- **The Protocol:**
  1. V picks a random $r \in \boldsymbol{F}$ and lets $\boldsymbol{x} = (r, r^2, \ldots, r^n)$.
  2. V computes $C \cdot \boldsymbol{x}$ and $(\mathrm{AB}) \cdot \boldsymbol{x}$, accepting iff they are equal.

Verifying Matrix Multiplication

- **The Protocol:**
  1. V picks a random $r \in \boldsymbol{F}$ and lets $\boldsymbol{x} = (r, r^2, \ldots, r^n)$.
  2. V computes $C \cdot \boldsymbol{x}$ and $(\mathrm{AB}) \cdot \boldsymbol{x}$, accepting iff they are equal.

- Runtime Analysis:
  - V's runtime dominated by computing 3 matrix-vector products, each of which takes $O(n^2)$ time.
    - $C \cdot \boldsymbol{x}$ is one matrix-vector multiplication.
    - $(\mathrm{AB}) \cdot \boldsymbol{x} = \mathrm{A} \cdot (\mathrm{B} \cdot \boldsymbol{x})$ takes two matrix-vector multiplications.

Correctness Analysis

- Claim 1: If $C = \mathrm{A} \cdot \mathrm{B}$ then V accepts with probability 1.
- Claim 2: If $C \neq \mathrm{A} \cdot \mathrm{B}$, then V rejects with probability at least

$$1 - \frac{n}{|\boldsymbol{F}|} \geq 1 - 1/n.$$

Correctness Analysis

- Claim 1: If $C = \mathrm{A} \cdot \mathrm{B}$ then V accepts with probability 1.
- Claim 2: If $C \neq \mathrm{A} \cdot \mathrm{B}$, then V rejects with probability at least

$$1 - \frac{n}{|\boldsymbol{F}|} \geq 1 - 1/n.$$

  - Proof of Claim 2:
    - Recall that $\boldsymbol{x} = (r, r^2, \dots, r^n)$.
    - $(C \cdot \boldsymbol{x})_i = \sum_{j=1}^{n} C_{ij} r^j$ is the Reed-Solomon fingerprint at r of the ith row of C.

Correctness Analysis

- Claim 1: If $C = \mathrm{A} \cdot \mathrm{B}$ then V accepts with probability 1.
- Claim 2: If $C \neq \mathrm{A} \cdot \mathrm{B}$, then V rejects with probability at least

$$1 - \frac{n}{|\boldsymbol{F}|} \geq 1 - 1/n.$$

  - Proof of Claim 2:
    - Recall that $\boldsymbol{x} = (r, r^2, \ldots, r^n)$.
    - $(C \cdot \boldsymbol{x})_i = \sum_{j=1}^{n} C_{ij} r^j$ is the Reed-Solomon fingerprint at r of the ith row of C.
    - Similarly, $((\mathrm{AB}) \cdot \boldsymbol{x})_i$ is the Reed-Solomon fingerprint at r of the ith row of AB.

Correctness Analysis

- Claim 1: If $C = \mathrm{A} \cdot \mathrm{B}$ then V accepts with probability 1.
- Claim 2: If $C \neq \mathrm{A} \cdot \mathrm{B}$, then V rejects with probability at least
$$1 - \frac{n}{|\boldsymbol{F}|} \geq 1 - 1/n.$$
  - Proof of Claim 2:
    - Recall that $\boldsymbol{x} = (r, r^2, \ldots, r^n)$.
    - $(C \cdot \boldsymbol{x})_i = \sum_{j=1}^{n} C_{ij} r^j$ is the Reed-Solomon fingerprint at r of the ith row of C.
    - Similarly, $((\mathrm{AB}) \cdot \boldsymbol{x})_i$ is the Reed-Solomon fingerprint at r of the ith row of AB.
    - So if even one row of C does not equal the corresponding row of AB, the fingerprints for that row will differ with probability at least $1 - 1/n$, causing V to reject.

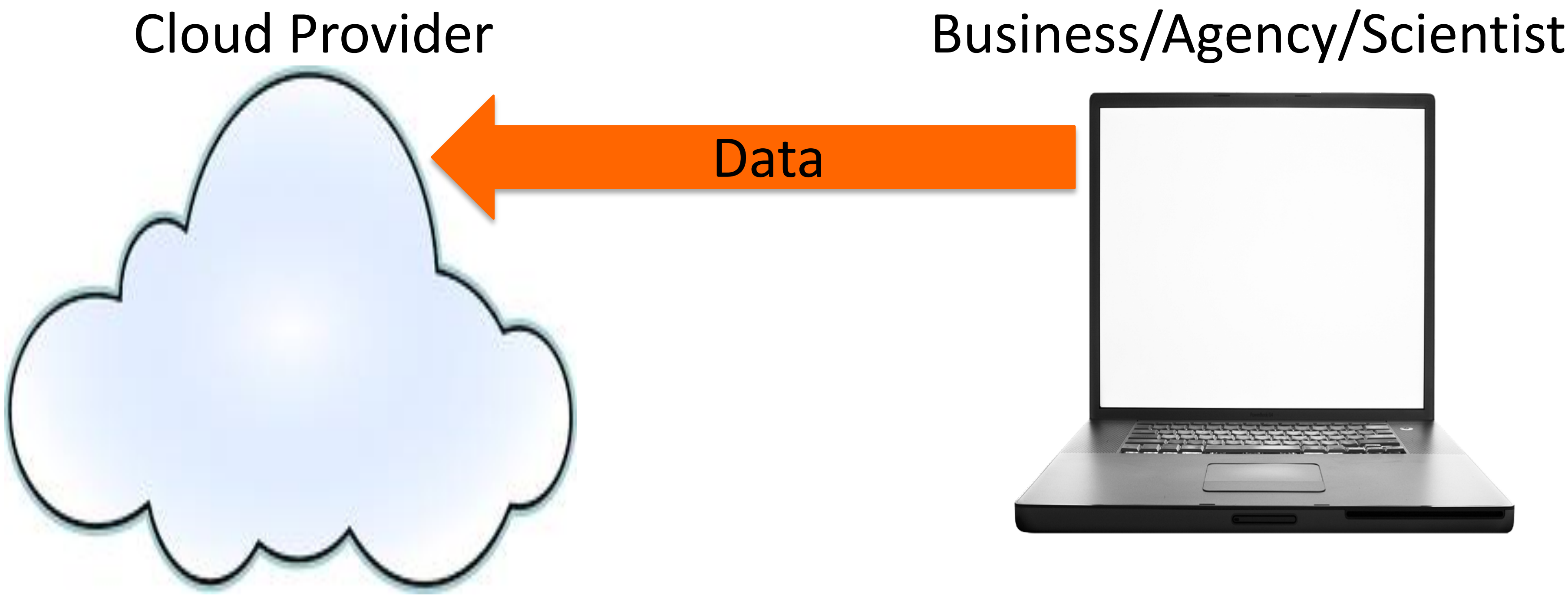
Interactive Proofs: Motivation and Model

Interactive Proofs

Cloud Provider

Business/Agency/Scientist

Interactive Proofs

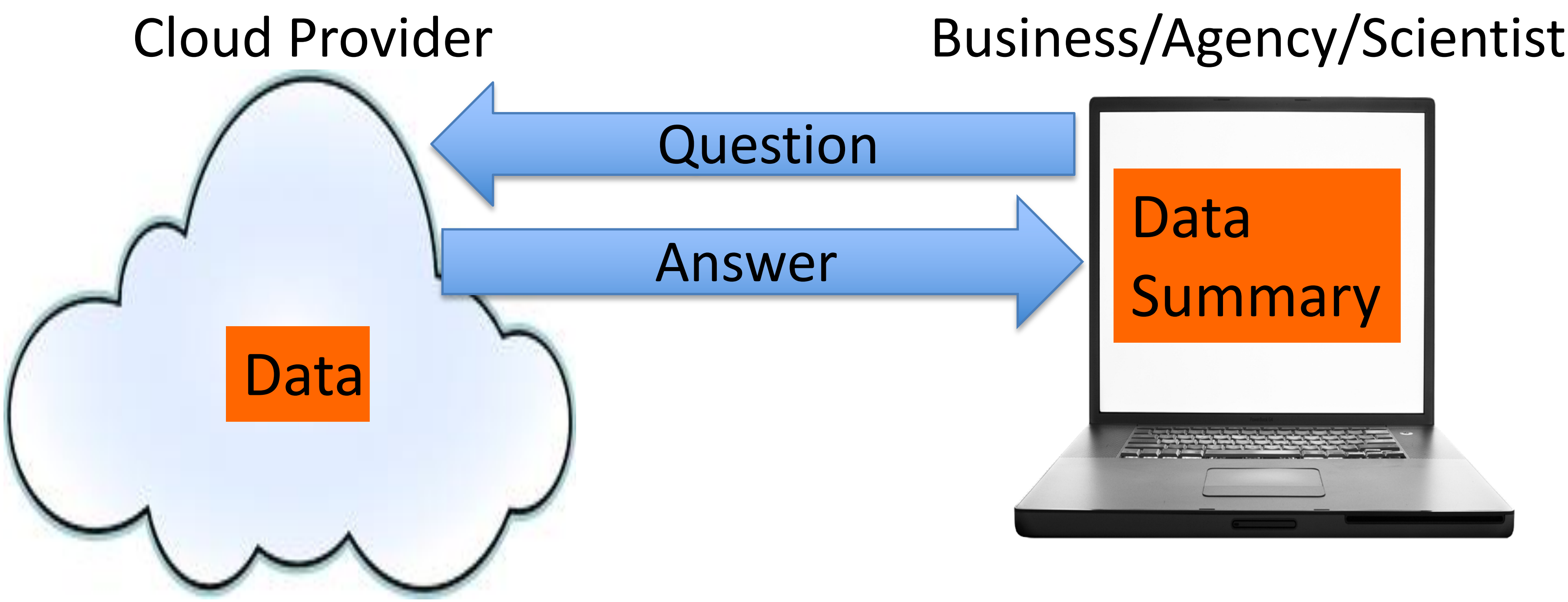


Interactive Proofs

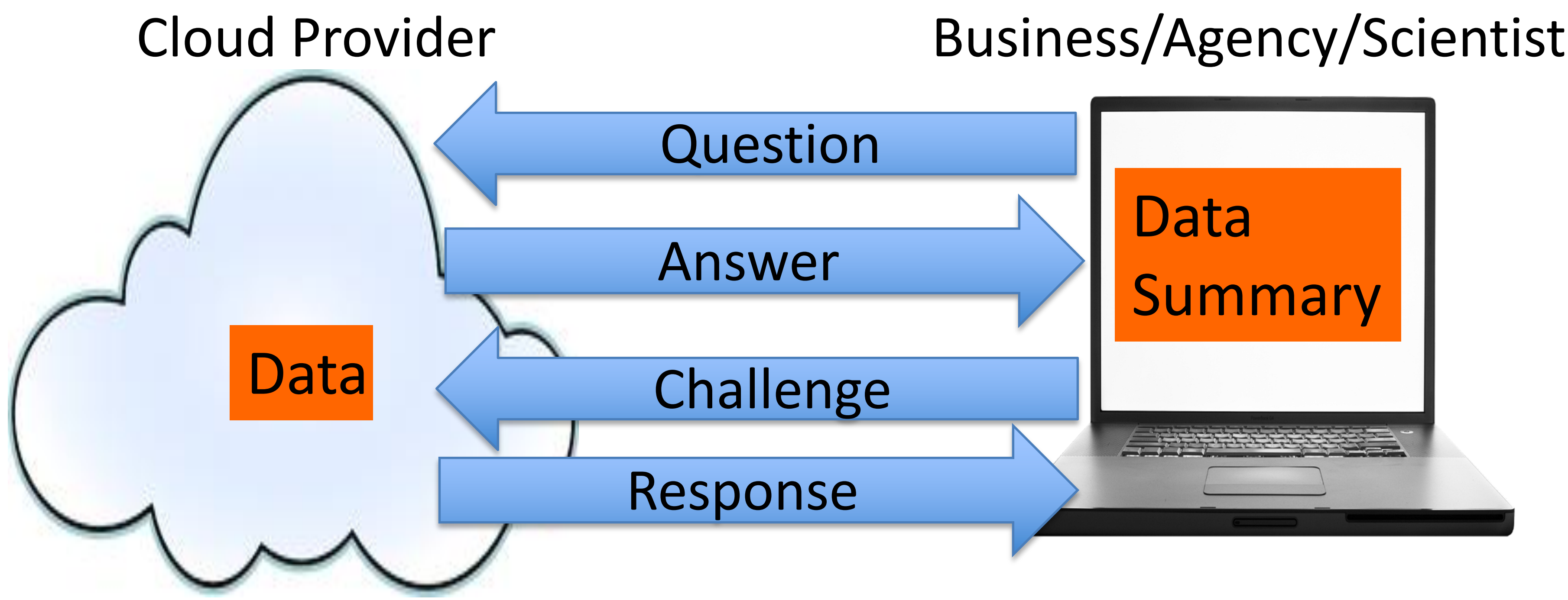
Cloud Provider

Business/Agency/Scientist

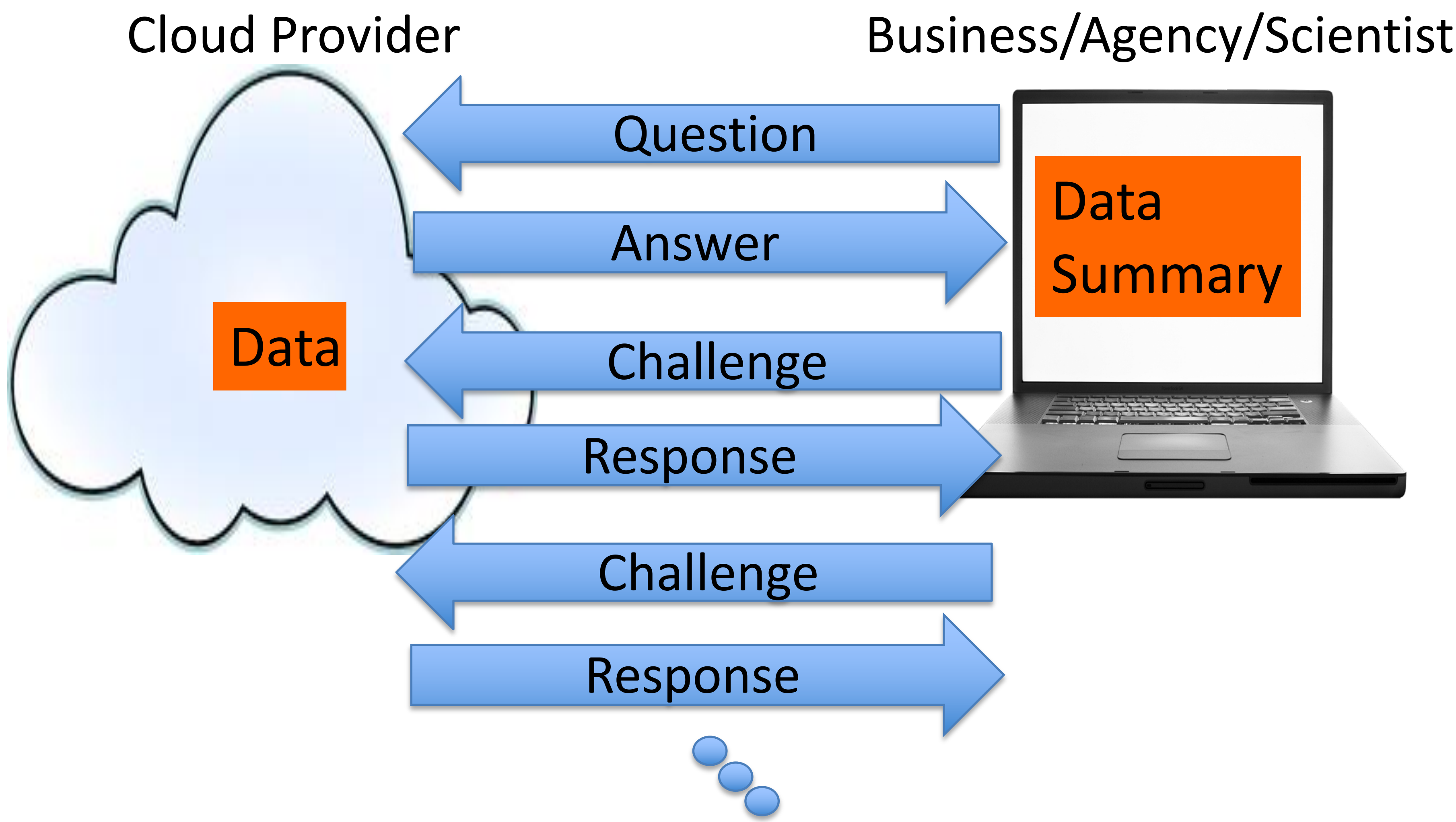
Interactive Proofs



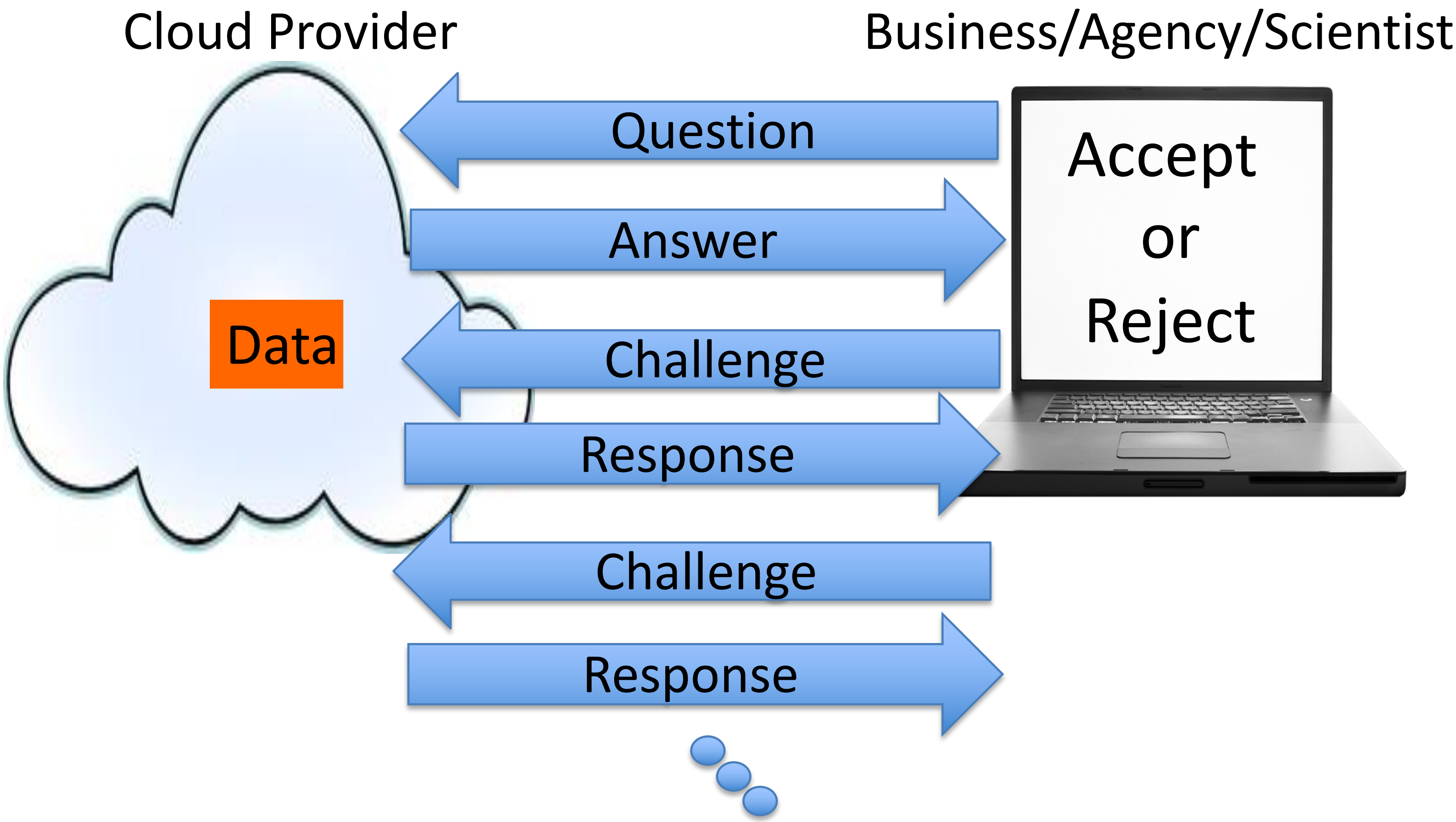
Interactive Proofs



Interactive Proofs



Interactive Proofs



Interactive Proofs

- Prover P and Verifier V.
- P solves problem, tells V the answer.
  - Then P and V have a conversation.
  - P's goal: convince V the answer is correct.
- Requirements:
  - 1. Completeness: an honest P can convince V to accept.
  - 2. Soundness: V will catch a lying P with high probability.

Interactive Proofs

- Prover P and Verifier V.

- P solves problem, tells V the answer.
  - Then P and V have a conversation.
  - P's goal: convince V the answer is correct.

- Requirements:
  - 1. Completeness: an honest P can convince V to accept.
  - 2. Soundness: V will catch a lying P with high probability.
    - This must hold even if P is computationally unbounded and trying to trick V into accepting the incorrect answer.

Interactive Proof Techniques: Preliminaries

Schwartz-Zippel Lemma

- Recall **FACT:** Let $p \neq q$ be univariate polynomials of degree at most d. Then $\Pr_{r \in \boldsymbol{F}}[p(r) = q(r)] \leq \frac{d}{|\boldsymbol{F}|}$.

Schwartz-Zippel Lemma

- Recall **FACT:** Let $p \neq q$ be univariate polynomials of degree at most d. Then $\Pr_{r \in \boldsymbol{F}}[p(r) = q(r)] \leq \frac{d}{|\boldsymbol{F}|}$.
- The **Schwartz-Zippel lemma** is a multivariate generalization:
  - Let $p \neq q$ be ℓ-variate polynomials of total degree at most d. Then $\Pr_{r \in \boldsymbol{F}^{\ell}}[p(r) = q(r)] \leq \frac{d}{|\boldsymbol{F}|}$.

Schwartz-Zippel Lemma

- Recall **FACT:** Let $p \neq q$ be univariate polynomials of degree at most d. Then $\Pr_{r \in \boldsymbol{F}}[p(r) = q(r)] \leq \frac{d}{|\boldsymbol{F}|}$.
- The **Schwartz-Zippel lemma** is a multivariate generalization:
  - Let $p \neq q$ be ℓ-variate polynomials of total degree at most d. Then $\Pr_{r \in \boldsymbol{F}^{\ell}}[p(r) = q(r)] \leq \frac{d}{|\boldsymbol{F}|}$.
  - "Total degree" refers to the maximum sum of degrees of all variables in any term. E.g., $x_1^2 x_2 + x_1 x_2$ has total degree 3.

Low-Degree and Multilinear Extensions

- Definition [**Extensions**]. Given a function $f: \{0,1\}^\ell \to \boldsymbol{F}$, a ℓ-variate polynomial g over **F** is said to **extend** f if $f(x) = g(x)$ for all $x \in \{0,1\}^\ell$.

- Definition [**Multilinear Extensions].** Any function $f: \{0,1\}^\ell \to \boldsymbol{F}$ has a **unique** multilinear extension (MLE), denoted $\tilde{f}$.

Low-Degree and Multilinear Extensions

- Definition [**Extensions**]. Given a function $f: \{0,1\}^{\ell} \to \boldsymbol{F}$, a ℓ-variate polynomial g over **F** is said to **extend** f if $f(x) = g(x)$ for all $x \in \{0,1\}^{\ell}$.

- Definition [**Multilinear Extensions].** Any function $f: \{0,1\}^{\ell} \to \boldsymbol{F}$ has a **unique** multilinear extension (MLE), denoted $\tilde{f}$.
  - Multilinear means the polynomial has degree at most 1 in each variable.
  - $(1 - x_1)(1 - x_2)$ is multilinear, $x_1^2 x_2$ is not.

$$f : \{0,1\}^2 \to \mathbf{F}$$

1	2
8	10

$$\tilde{f} : \mathbf{F}^2 \rightarrow \mathbf{F}$$

1	2	3	4	5	6
8	10	12	14	16	18
15	18	21	24	27	30
22	26	30	34	38	42
29	34	39	44	49	56
36	42	48	54	60	68

$$\tilde{f}(x_1, x_2) = (1 - x_1)(1 - x_2) + 2(1 - x_1)x_2 + 8x_1(1 - x_2) + 10x_1x_2$$

1	2	3	4	5	6
8	10	12	14	16	18
15	18	21	24	27	30
22	26	30	34	38	42
29	34	39	44	49	56
36	42	48	54	60	68

Can check:

$\tilde{f}(0,0) = 1$

$\tilde{f}(0,1) = 2$

$\tilde{f}(1,0) = 8$

$\tilde{f}(1,1) = 10$

Another (non-multilinear) extension of f:

$$g(x_1, x_2) = -x_1^2 + x_1 x_2 + 8\, x_1 + \; x_2 + 1$$

1	2	3	4	5	6	
8	10	12	14	16	18	
13	16	19	22	25	28	…
16	20	24	28	32	36	
17	22	27	32	37	42	
16	22	28	34	40	44	
		⋮				

Can check:

$g(0,0) = 1$

$g(0,1) = 2$

$g(1,0) = 8$

$g(1,1) = 10$