Data provenance in a distributed calculus

Motivation

- (Meta)data is almost entirely neglected in the process calculi literature
- Track data provenance both for its important applications and as an challenging exercise in modelling (meta)data. We aim at simplicity:
  - data annotations representing provenance
  - structure, interpretation and management of provenance information
  - provenance tracking
- Provenance-based security (aspects: trust + data confidentiality and privacy)
  - Example: photography competition
- The overall ambition is to underpin practical development, like trust-policy languages and protocols, and provenance-middleware

Provenance model

Annotated data

$v : \kappa$

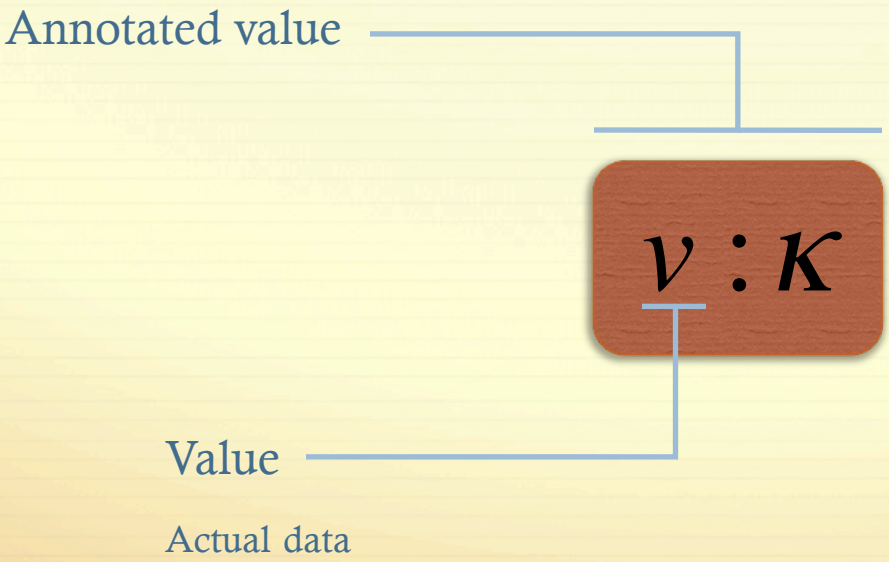
Provenance model

Annotated data

Annotated value

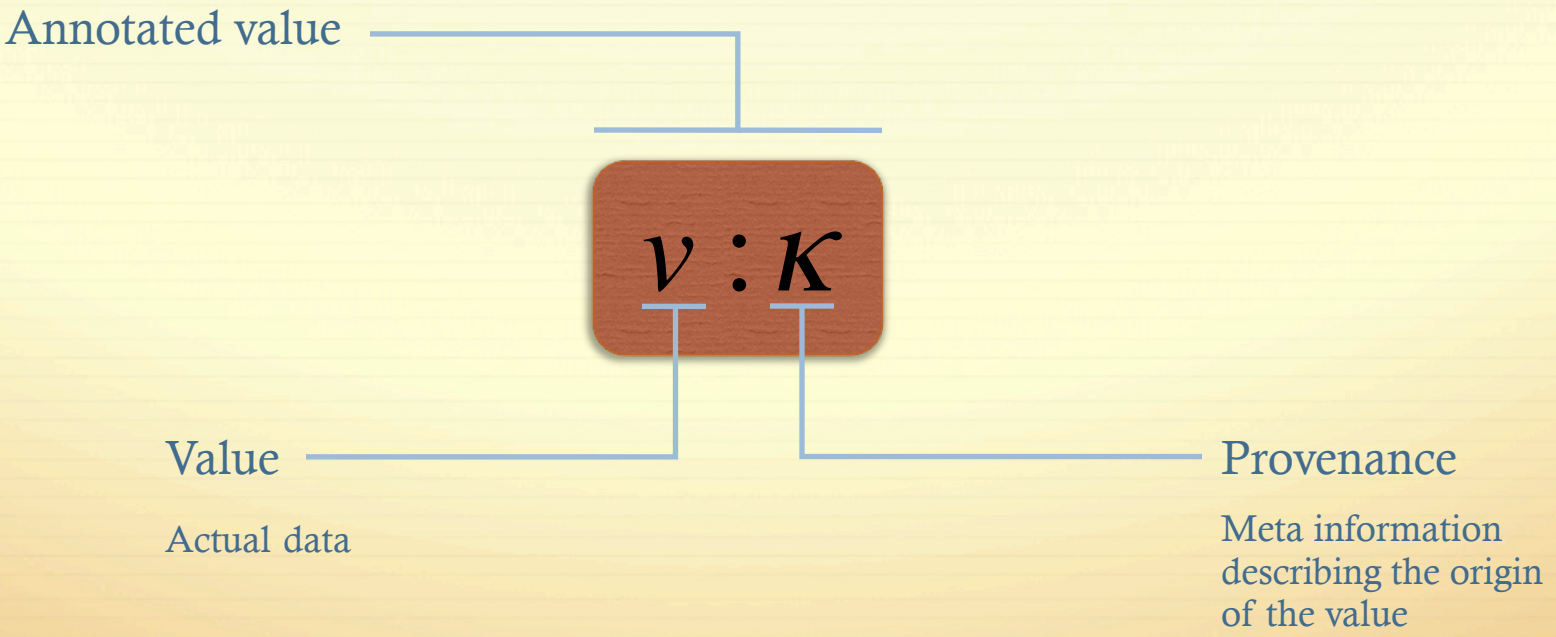
Provenance model

Annotated data



Provenance model

Annotated data



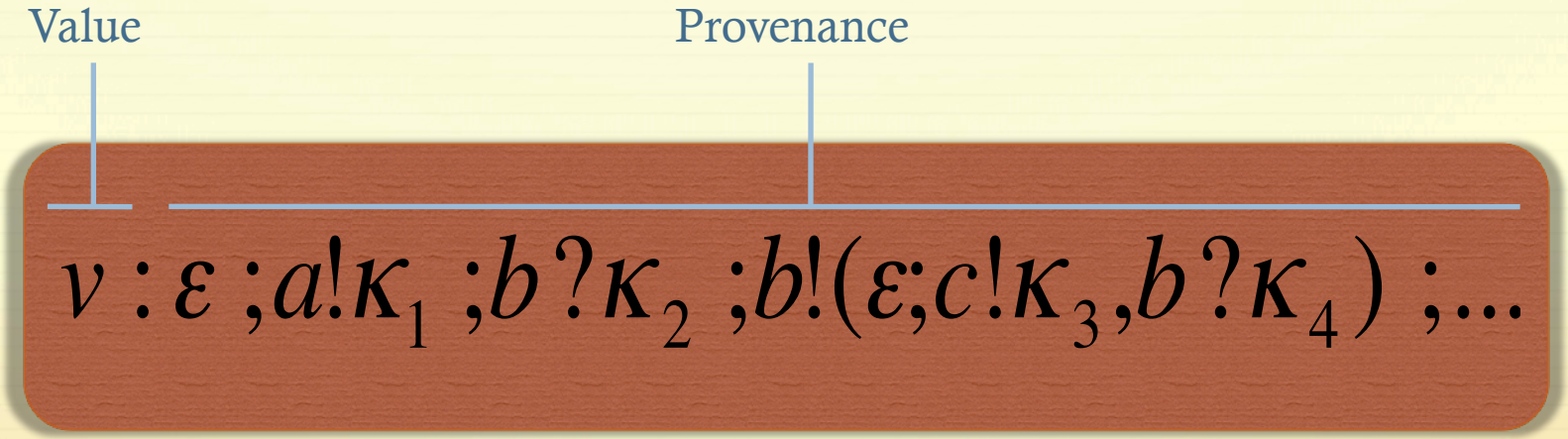
Provenance model

Structure and interpretation of provenance

$$v : \varepsilon \; ; a!\kappa_1 \; ; b?\kappa_2 \; ; b!(\varepsilon ; c!\kappa_3 , b?\kappa_4) \; ; \ldots$$

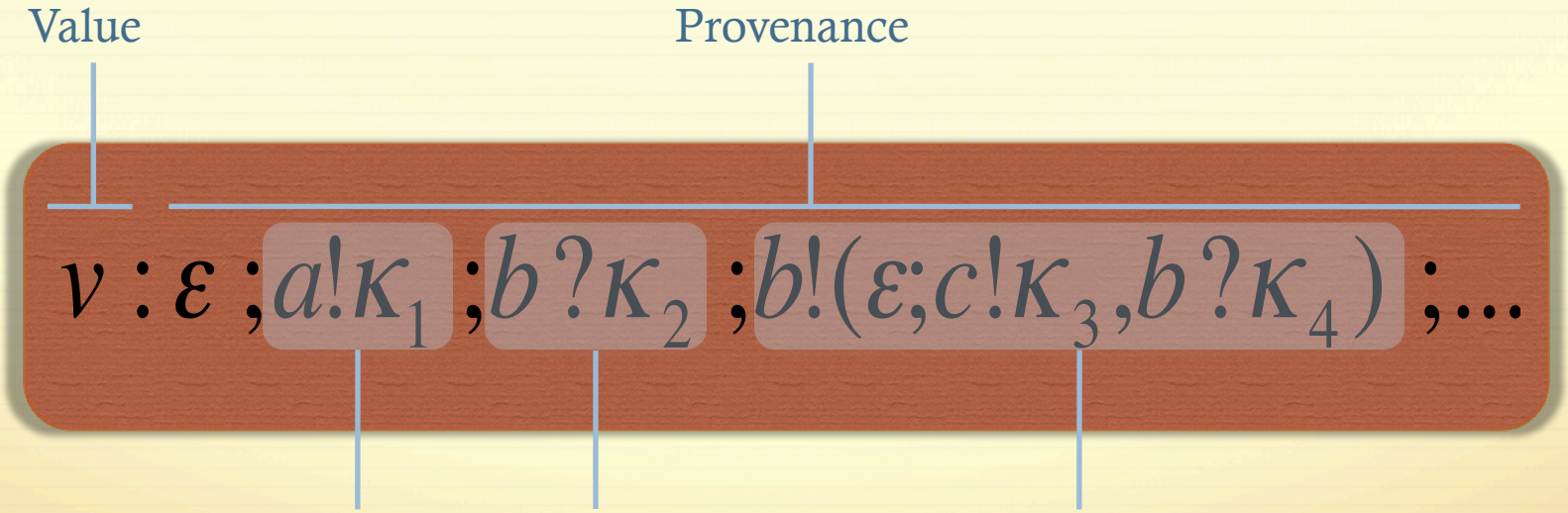
Provenance model

Structure and interpretation of provenance



Provenance model

Structure and interpretation of provenance



Provenance model

Structure and interpretation of provenance

ε (empty provenance)
denotes value v originated here

Provenance model

Structure and interpretation of provenance

ε (empty provenance)
denotes value v originated here

It was sent by a on a channel with provenance κ_1

Provenance model

Structure and interpretation of provenance

ε (empty provenance)
denotes value v originated here

It was sent by a on a channel with provenance κ_1

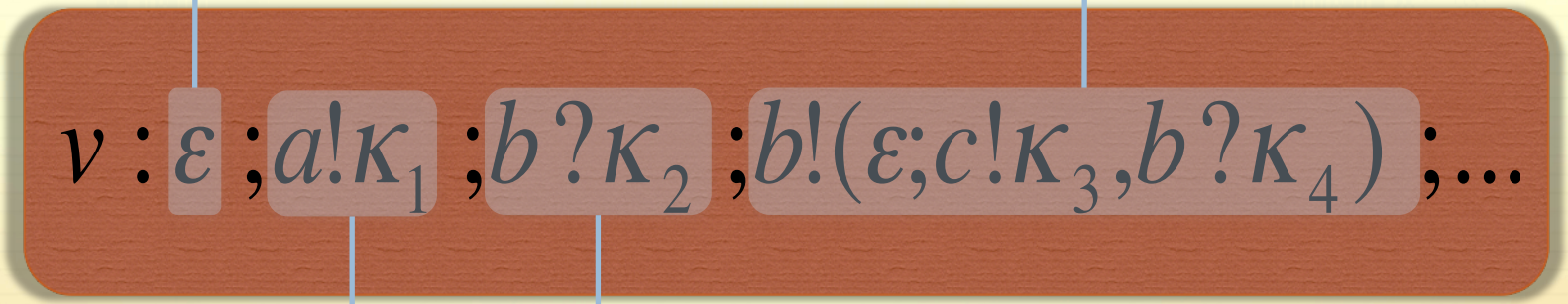
Was then received by b on a channel with provenance κ_2

Provenance model

Structure and interpretation of provenance

ε (empty provenance) denotes value v originated here

And then sent by b on a channel that b received from c…



It was sent by a on a channel with provenance κ_1

Was then received by b on a channel with provenance κ_2

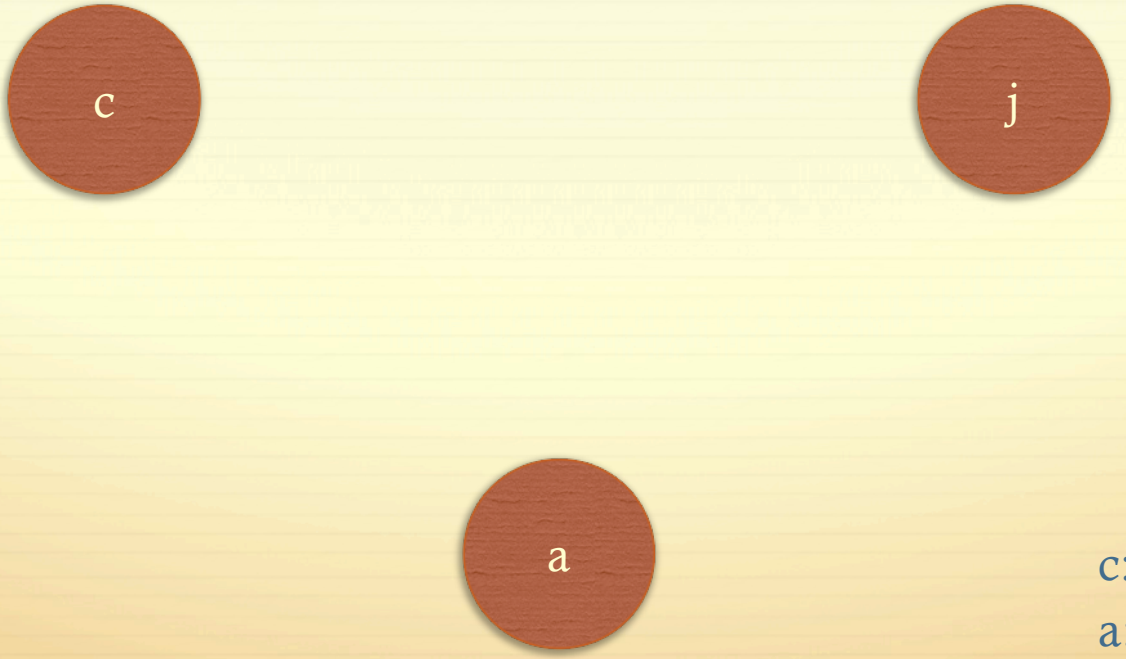
Confidentiality in provenance systems

- Data may be public, yet its provenance confidential, or vice versa
- Principals who may access data are not necessarily the same as those who may access its provenance
- In general, fine grained access control over provenance "histories" is needed as different parts of it have different sensitivity

Security requirements of data ≠ Security requirements of its provenance

Hiding provenance trees

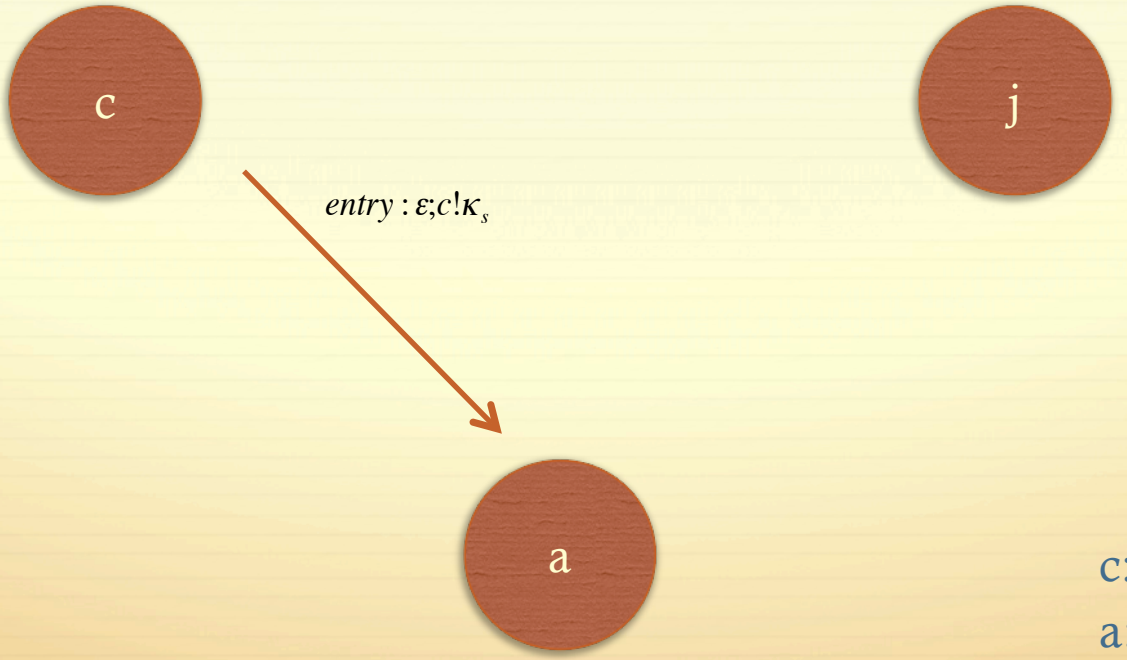
Example: photography competition



c: Contestant
a: Administrator
j: Judge

Hiding provenance trees

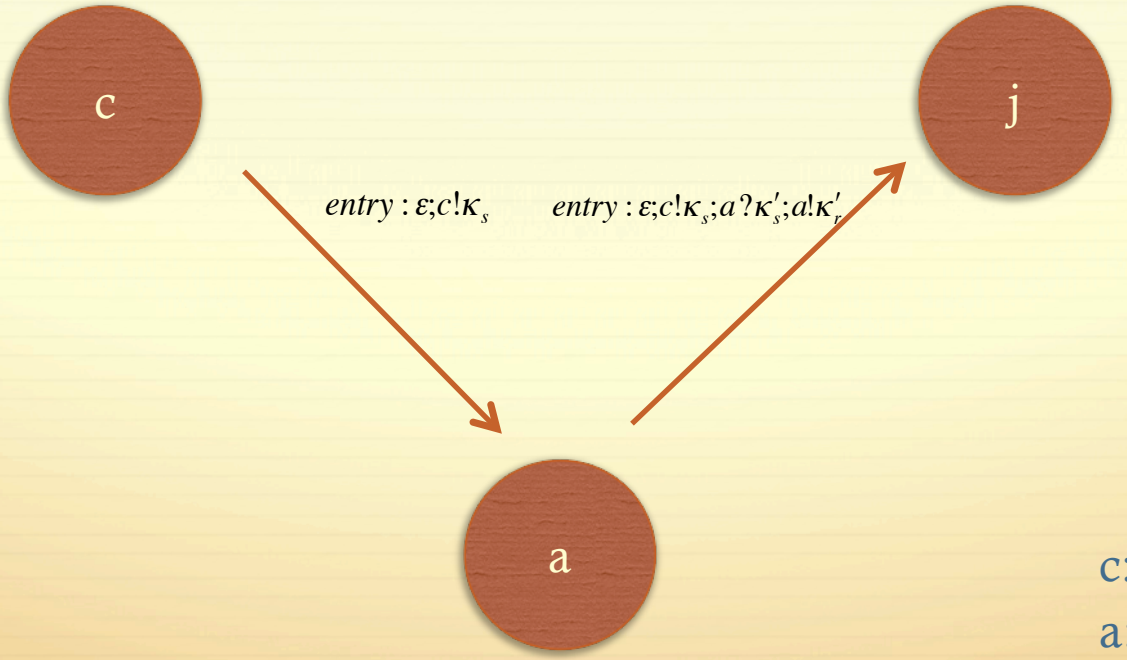
Example: photography competition



c: Contestant
a: Administrator
j: Judge

Hiding provenance trees

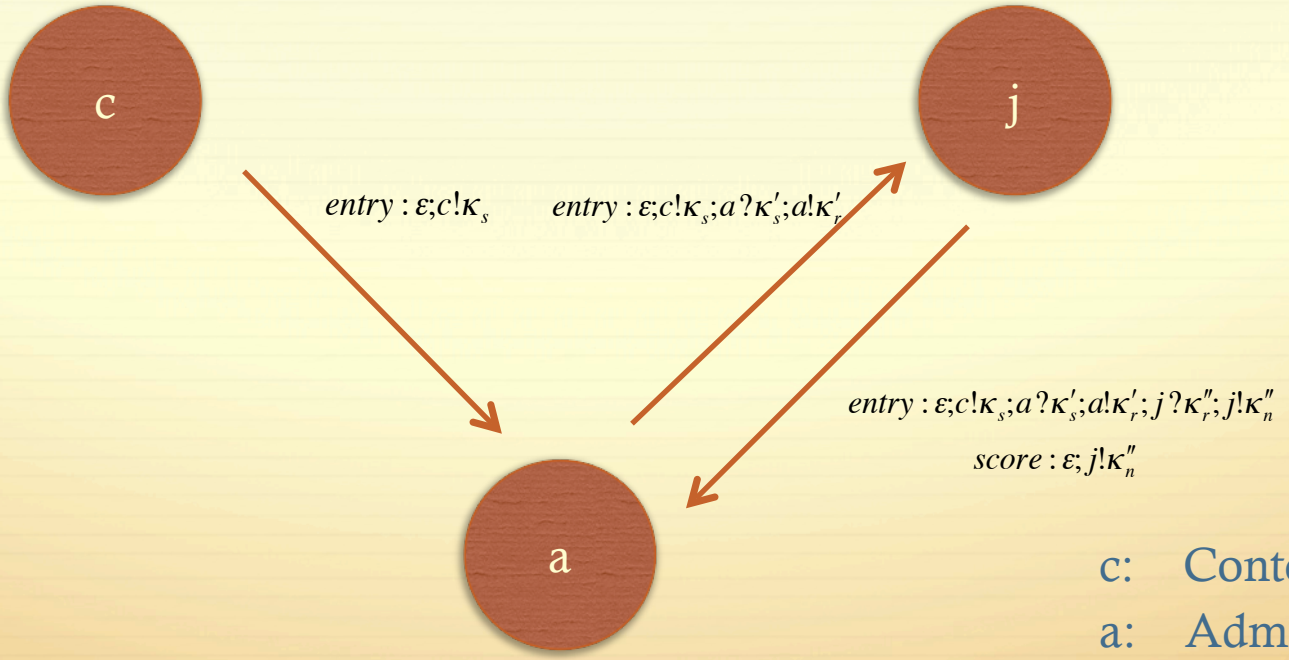
Example: photography competition



c: Contestant
a: Administrator
j: Judge

Hiding provenance trees

Example: photography competition



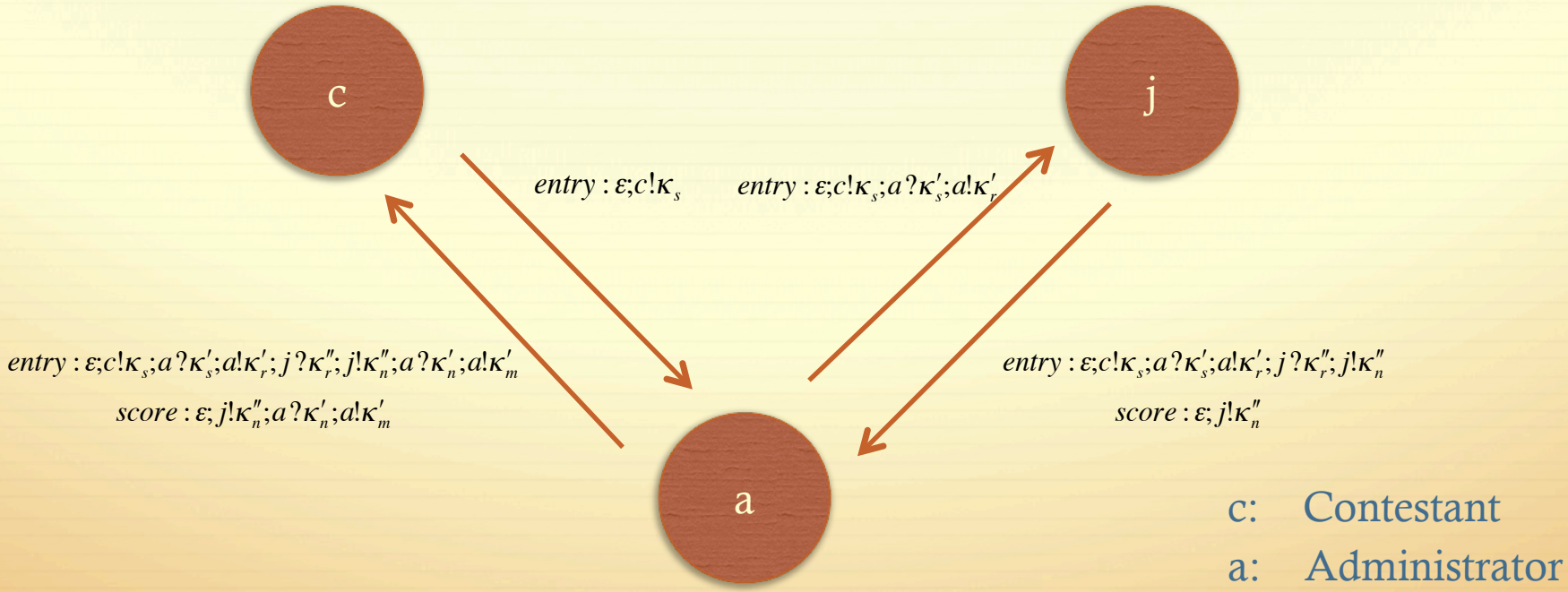
c: Contestant

a: Administrator

j: Judge

Hiding provenance trees

Example: photography competition



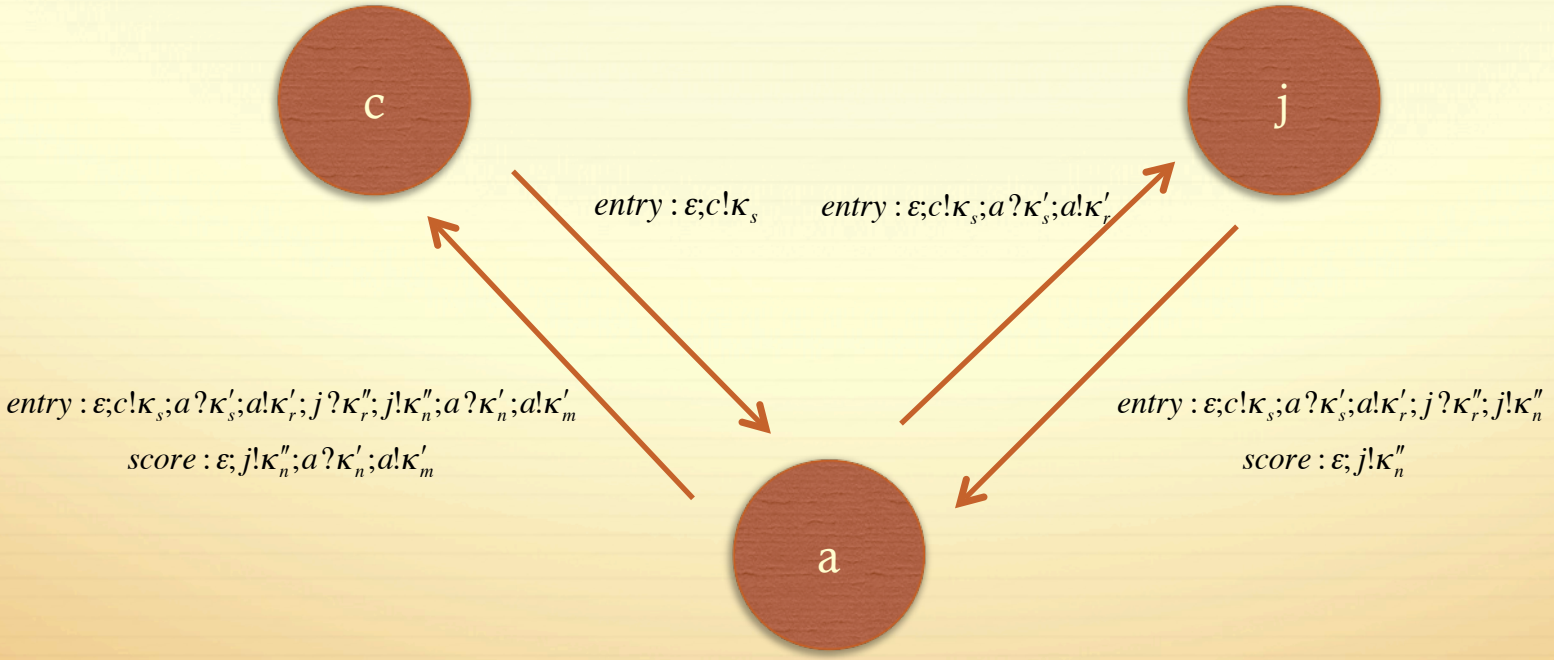
c: Contestant

a: Administrator

j: Judge

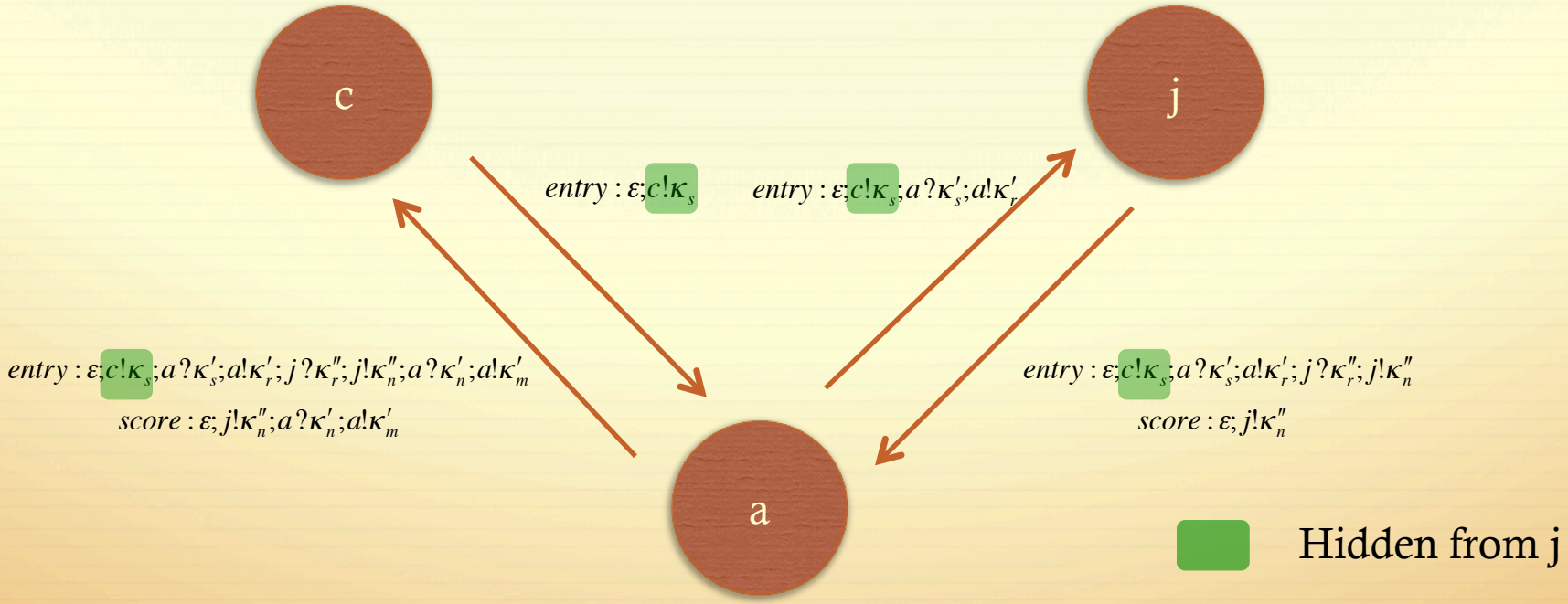
Hiding provenance trees

Example: photography competition



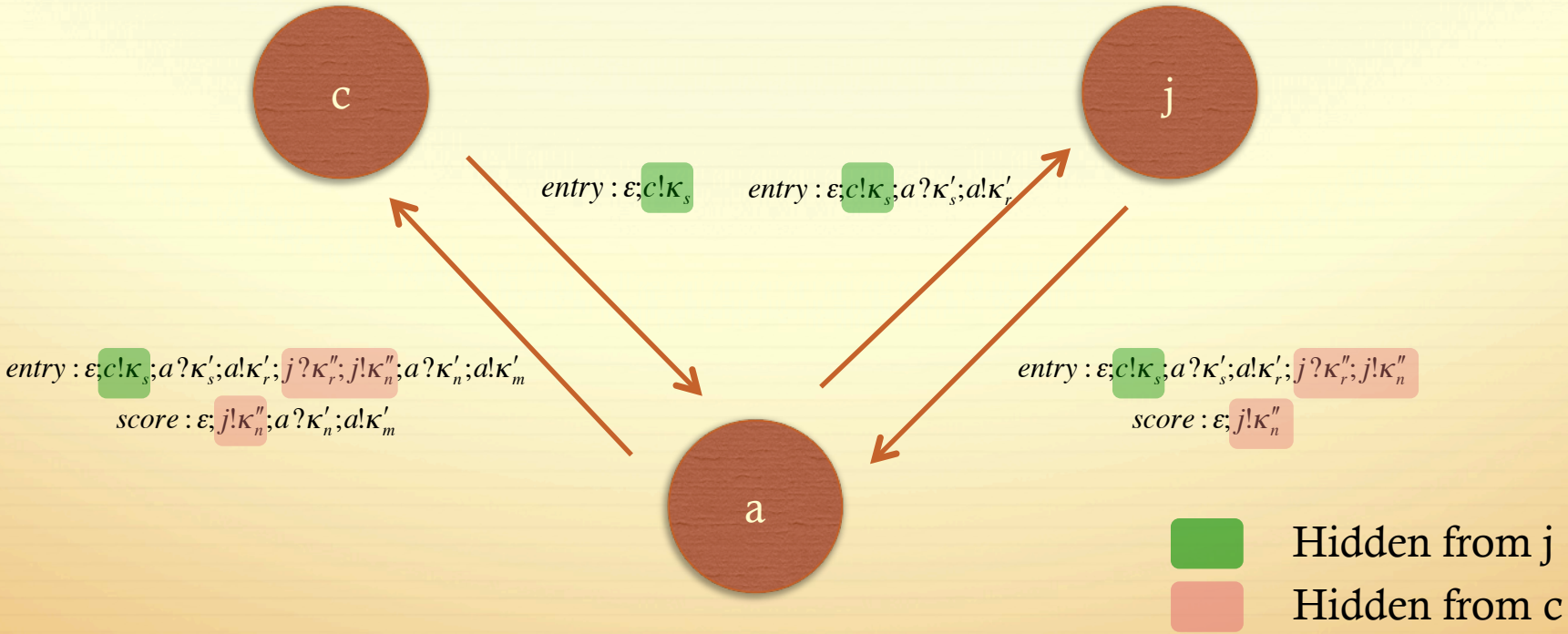
Hiding provenance trees

Example: photography competition



Hiding provenance trees

Example: photography competition



Confidentiality in provenance systems

a promising approach

- One value, multiple **views**
  - Different principals have different views of the same provenance list based on their privileges

$$entry : \varepsilon; c!\kappa_s; a?\kappa'_s; a!\kappa'_r; j?\kappa''_r; j!\kappa''_n; a?\kappa'_n; a!\kappa'_m$$

Confidentiality in provenance systems

a promising approach

- One value, multiple **views**
  - Different principals have different views of the same provenance list based on their privileges

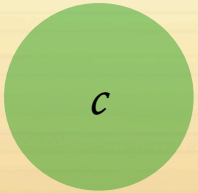
$$entry : \varepsilon; c!\kappa_s; a?\kappa'_s; a!\kappa'_r; j?\kappa''_r; j!\kappa''_n; a?\kappa'_n; a!\kappa'_m$$

Confidentiality in provenance systems

a promising approach

- One value, multiple **views**
  - Different principals have different views of the same provenance list based on their privileges

$$entry : \varepsilon; c!\kappa_s; a?\kappa'_s; a!\kappa'_r; j?\kappa''_r; j!\kappa''_n; a?\kappa'_n; a!\kappa'_m$$



Confidentiality in provenance systems

a promising approach

- One value, multiple **views**
  - Different principals have different views of the same provenance list based on their privileges

$$entry : \varepsilon; c!\kappa_s; a?\kappa'_s; a!\kappa'_r; j?\kappa''_r; j!\kappa''_n; a?\kappa'_n; a!\kappa'_m$$

Confidentiality in provenance systems

a promising approach

- One value, multiple **views**
  - Different principals have different views of the same provenance list based on their privileges

$$entry : \varepsilon; c!\kappa_s; a?\kappa'_s; a!\kappa'_r; j?\kappa''_r; j!\kappa''_n; a?\kappa'_n; a!\kappa'_m$$