The background of the slide is a collage of various vintage postage stamps and postmarks. Visible elements include a red circular postmark with the word 'PARAVION' and the number '02293', a yellow circular postmark with the word 'POSTAGE', a red rectangular postmark with the word 'MADRID', and a red circular postmark with the word 'COSTA RICA'. There are also various numbers and other markings scattered across the collage.

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 a 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



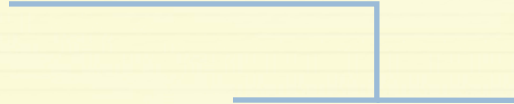
$v : K$

Provenance model

Annotated data



Annotated value



$v : K$

Provenance model

Annotated data



Annotated value

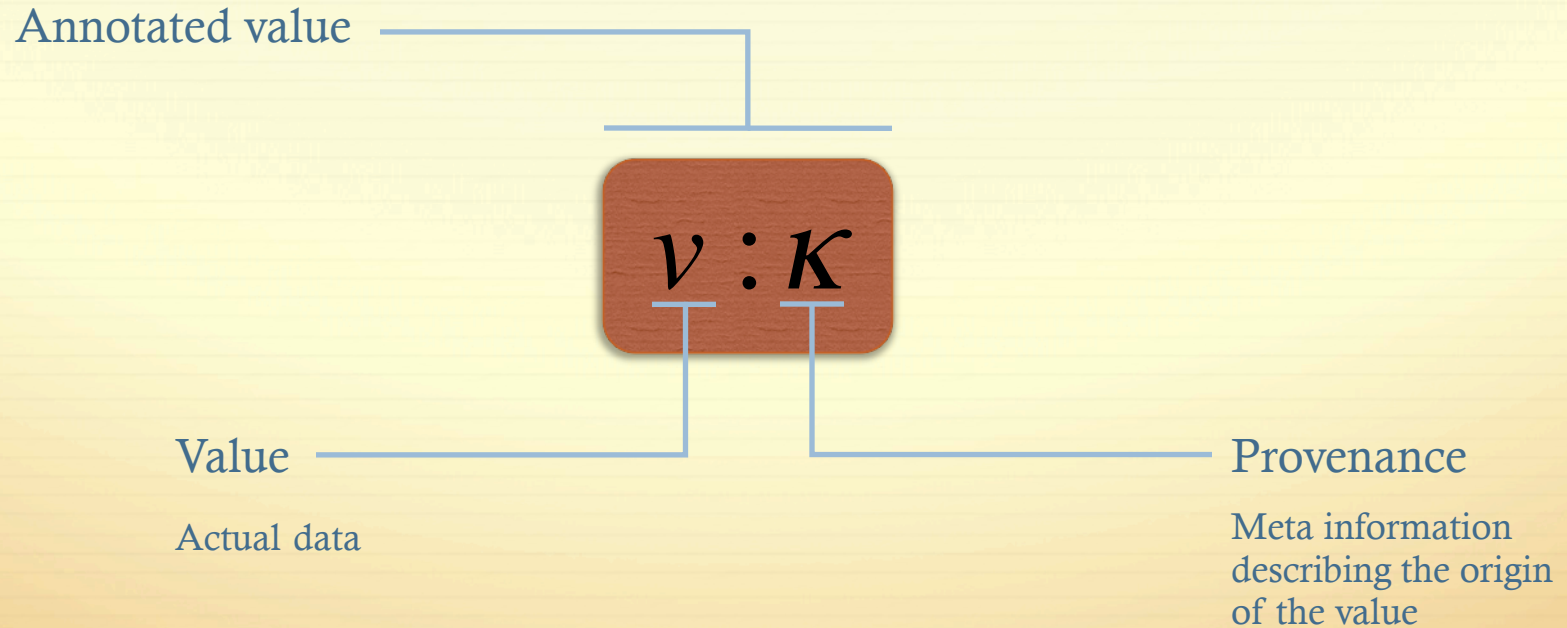


Value

Actual data

Provenance model

Annotated data



Provenance model

Structure and interpretation of provenance

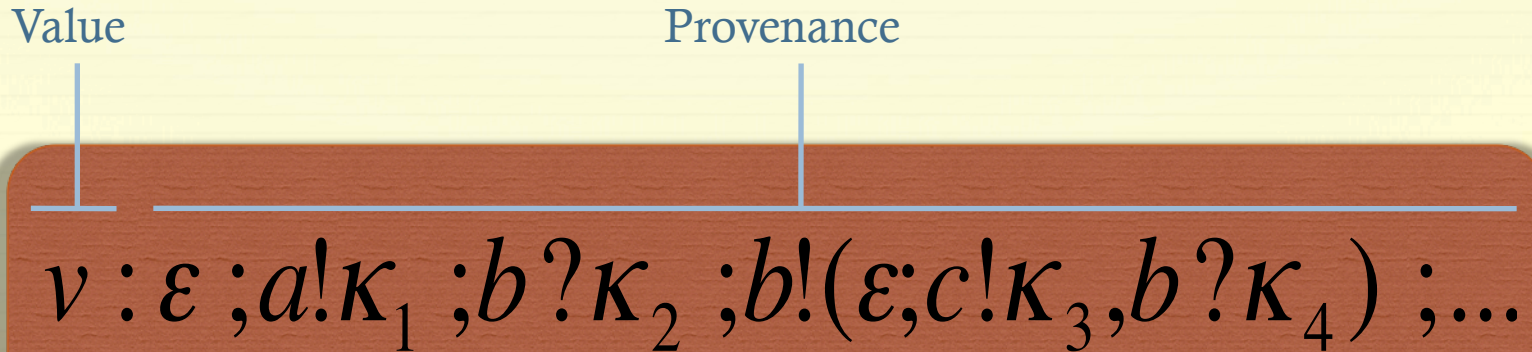

$$v : \varepsilon ; a! \kappa_1 ; b? \kappa_2 ; b!(\varepsilon ; c! \kappa_3 , b? \kappa_4) ; \dots$$

Provenance model

Structure and interpretation of provenance

Value

Provenance



The diagram illustrates the structure of a provenance model. It features a horizontal line with two vertical lines extending upwards from it. The left vertical line is labeled 'Value' and the right vertical line is labeled 'Provenance'. Below the horizontal line, the expression $v : \varepsilon ; a ! \kappa_1 ; b ? \kappa_2 ; b ! (\varepsilon ; c ! \kappa_3 , b ? \kappa_4) ; \dots$ is written. The entire diagram is enclosed in a rounded rectangle with a dark red background.

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

Provenance model

Structure and interpretation of provenance

Value

Provenance

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


“Operations” that were performed on the value. They record the principals that “influenced” the value and how.

Provenance model

Structure and interpretation of provenance

ε (empty provenance)

denotes value v originated here



The diagram shows a large, horizontal, rounded rectangle with a textured, reddish-brown background. On the left side of this rectangle, there is a small, light-colored rectangular box containing the text $v : \varepsilon$. A thin vertical line extends upwards from the top of this box, ending in a small diamond shape. The entire diagram is set against a light yellow background with faint, stylized text and graphics.

$v : \varepsilon$

Provenance model

Structure and interpretation of provenance

ε (empty provenance)
denotes value v originated here

$v : \varepsilon ; a ! \kappa_1$

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

$v : \varepsilon ; a! \kappa_1 ; b? \kappa_2$

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 $c...$

The diagram shows a provenance expression $v : \varepsilon ; a! \kappa_1 ; b? \kappa_2 ; b!(\varepsilon ; c! \kappa_3 , b? \kappa_4) ; \dots$ inside a large rounded rectangle. The expression is composed of several parts, each highlighted in a light purple box: v , ε , $a! \kappa_1$, $b? \kappa_2$, $b!(\varepsilon ; c! \kappa_3 , b? \kappa_4)$, and \dots . Blue lines connect these parts to explanatory text blocks. A line from ε points to the top-left text. A line from $a! \kappa_1$ points to the bottom-left text. A line from $b? \kappa_2$ points to the bottom-right text. A line from $b!(\varepsilon ; c! \kappa_3 , b? \kappa_4)$ points to the top-right text.

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

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

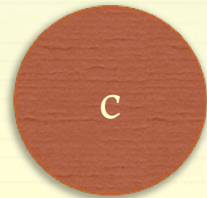
Security requirements of
data

≠

Security requirements of its
provenance

Hiding provenance trees

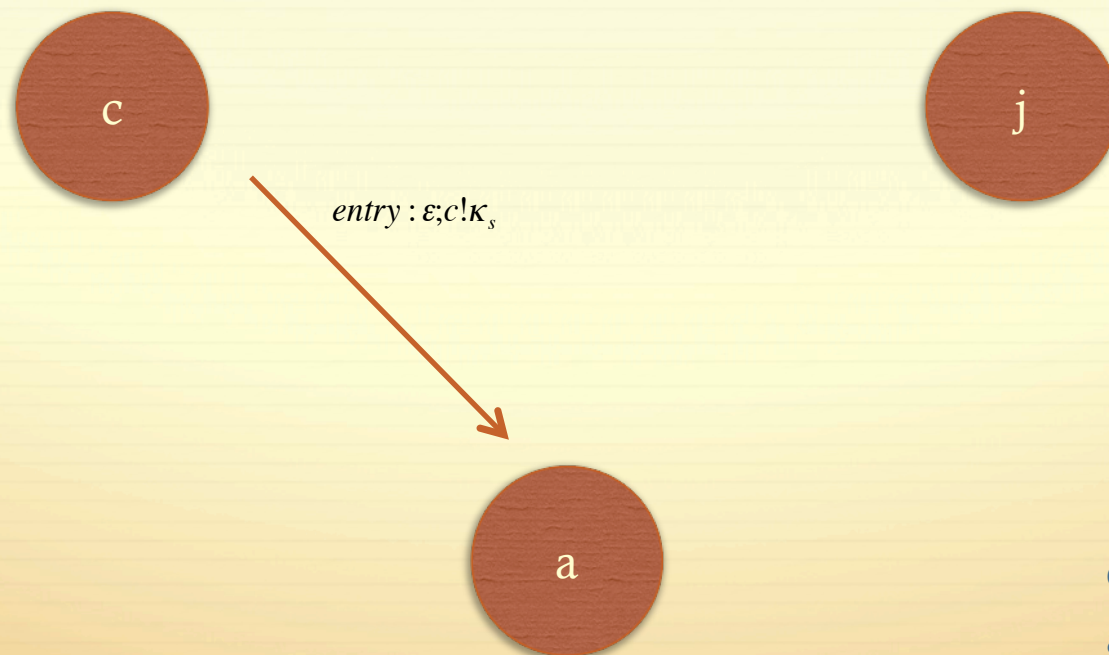
Example: photography competition



c: Contestant
a: Administrator
j: Judge

Hiding provenance trees

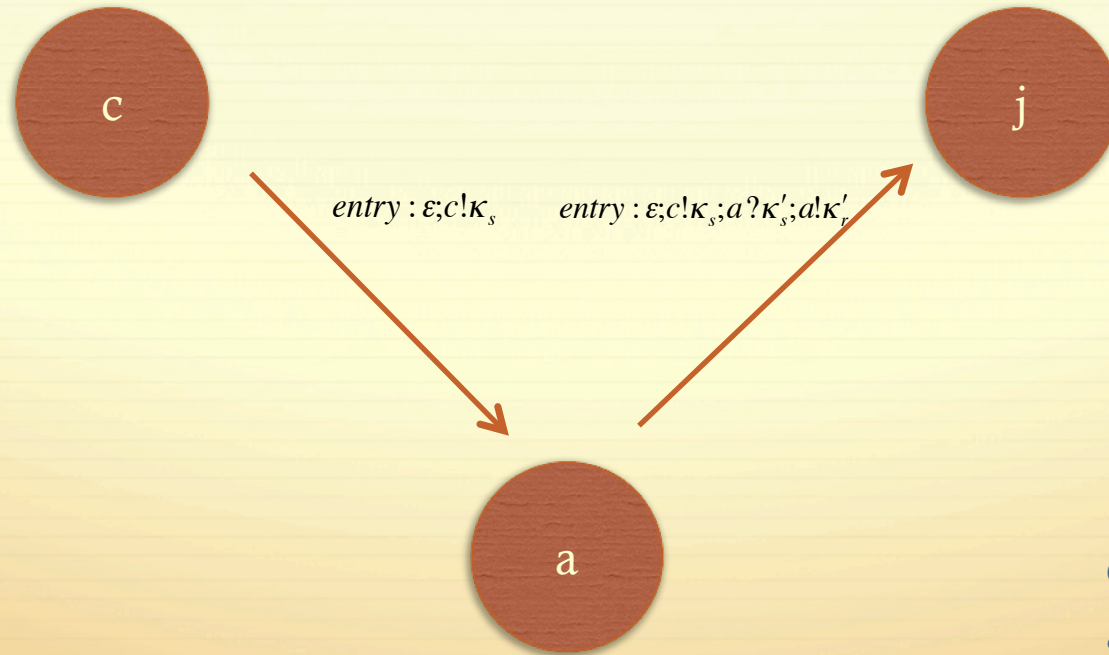
Example: photography competition



c: Contestant
a: Administrator
j: Judge

Hiding provenance trees

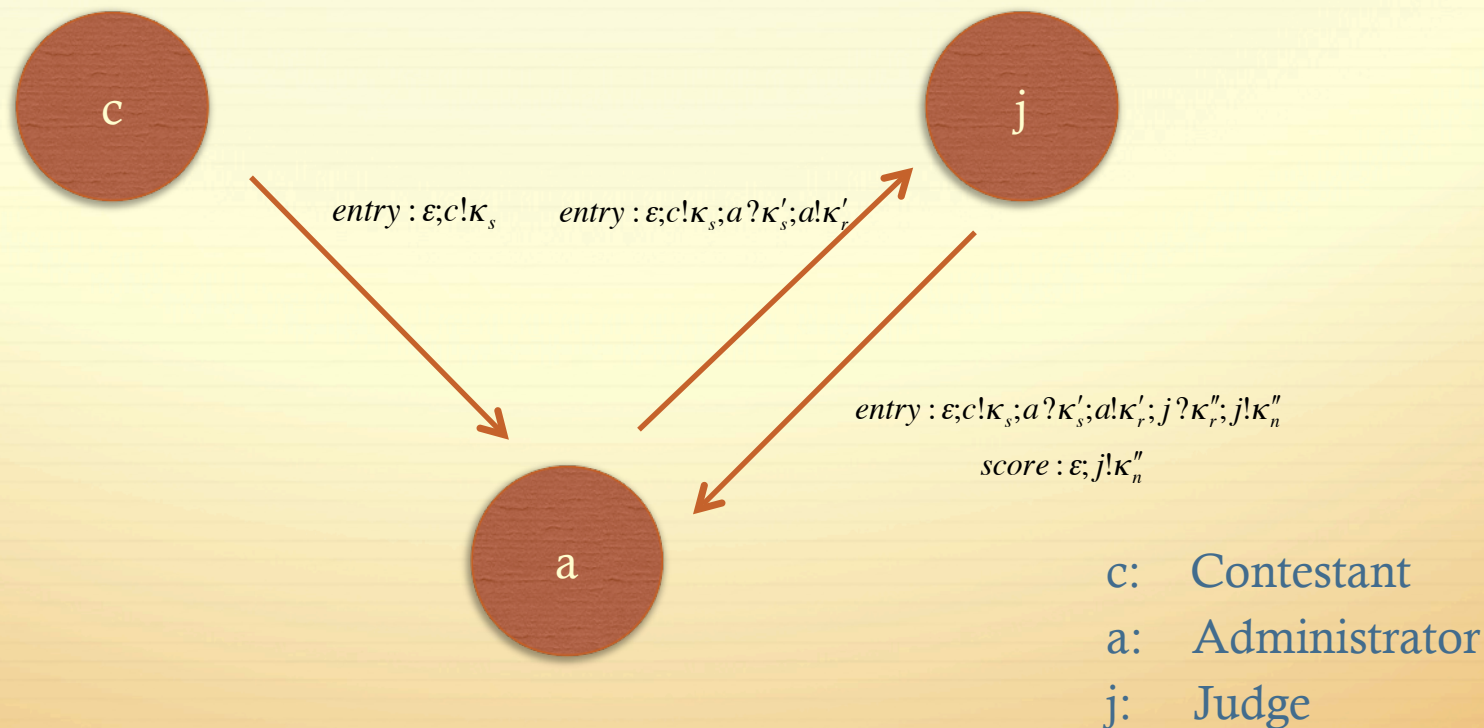
Example: photography competition



c: Contestant
a: Administrator
j: Judge

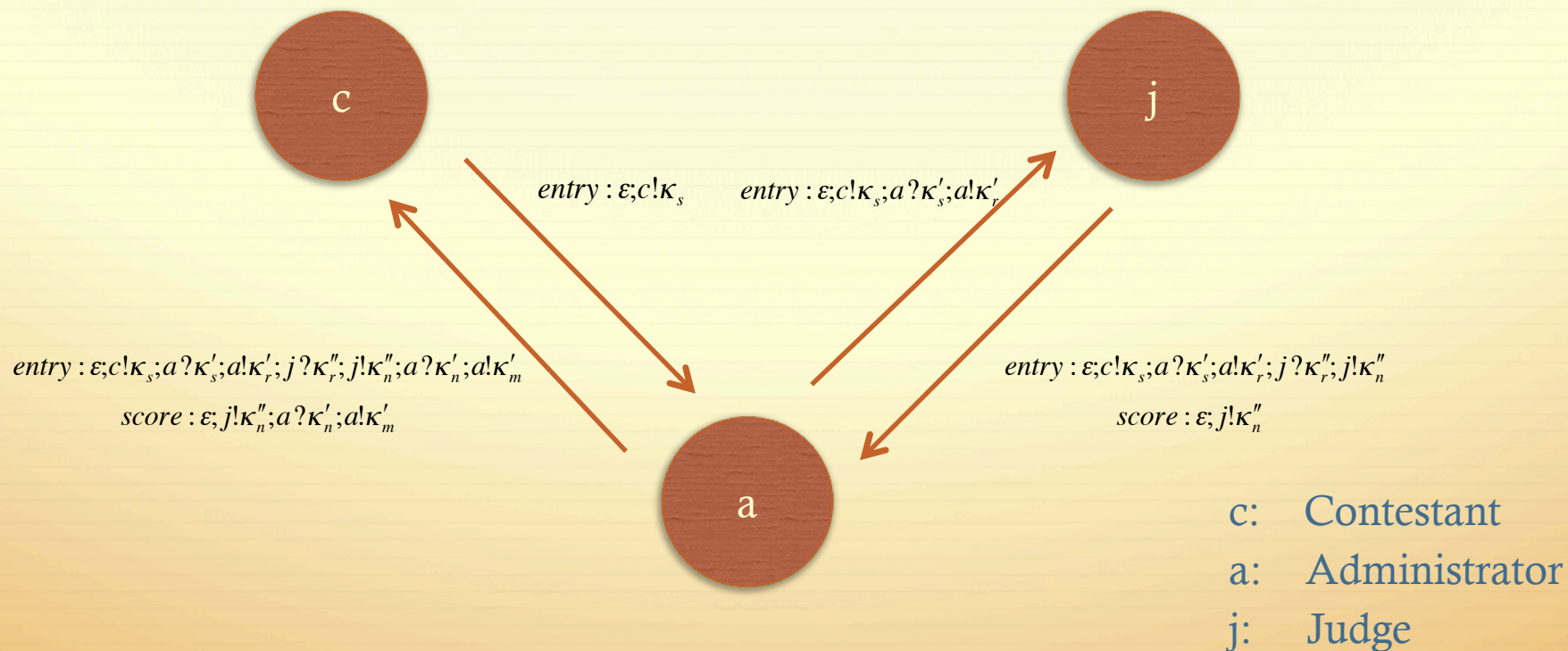
Hiding provenance trees

Example: photography competition



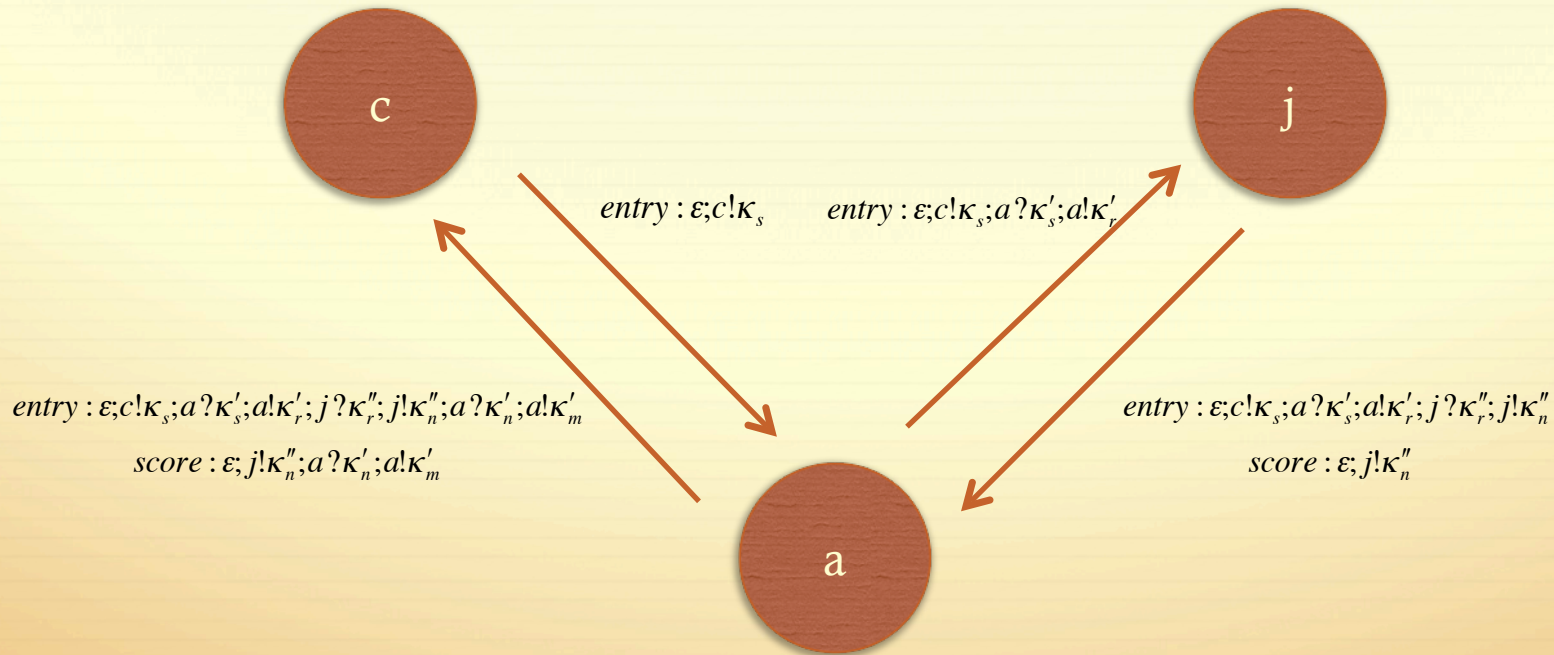
Hiding provenance trees

Example: photography competition



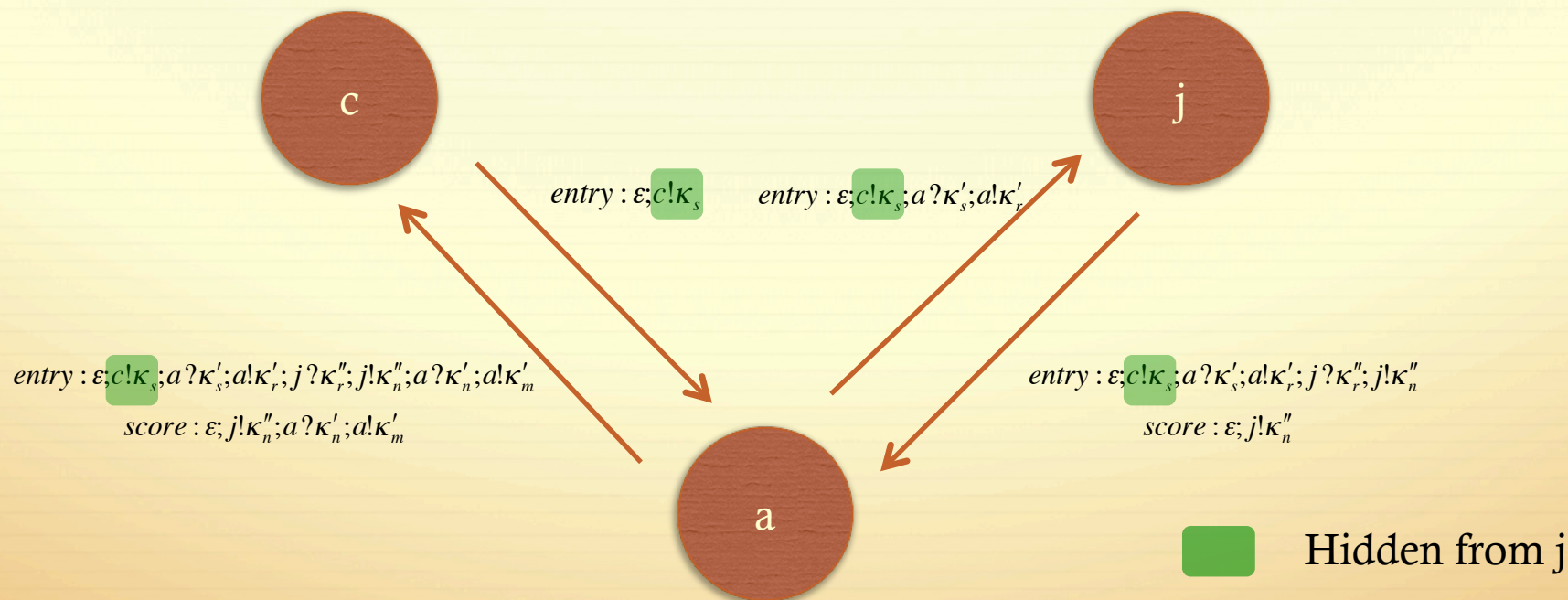
Hiding provenance trees

Example: photography competition



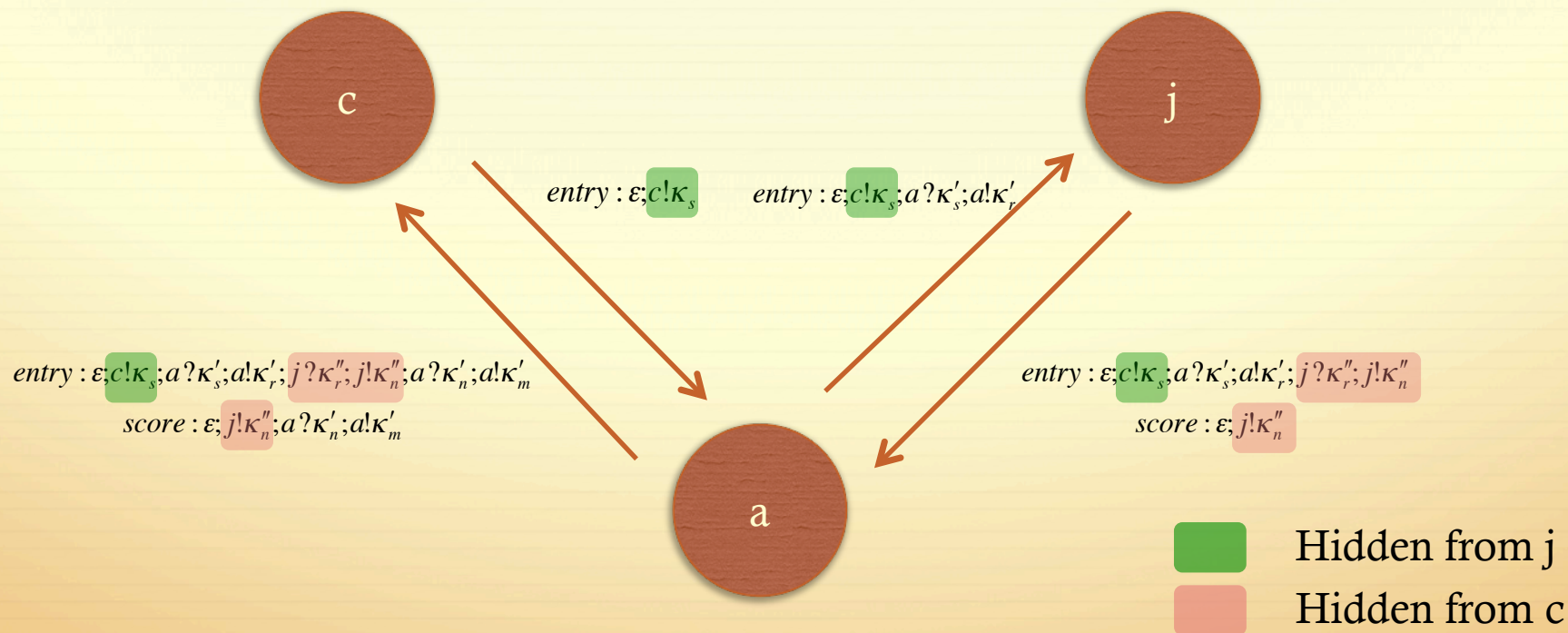
Hiding provenance trees

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



a

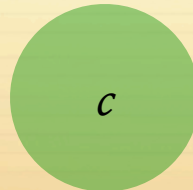
Confidentiality in provenance systems

a promising approach



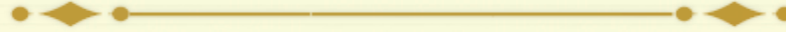
- ✧ One value, multiple **views**
- ✧ Different principals have different views of the same provenance list based on their privileges

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

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

j

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

