Relational Database Design

Module 7: Basic Normalization (Part 2)

Hugo Kornelis
hugo@perFact.info

pluralsight
hardcore developer training

Outline

- Normalization
- Functional dependencies
- Basic normal forms
  - First Normal Form (1NF)
  - Second Normal Form (2NF)
  - Third Normal Form (3NF)
- **Finding functional dependencies**

Second Normal Form

- **Requirements for Second Normal Form (2NF):**
  - Table must be in First Normal Form
  - Non-key attributes must not depend on subset of any candidate key
    - "Non-key attributes" – not part of *any* candidate key
    - "Subset of candidate key" – requires at least one composite candidate key

No composite keys

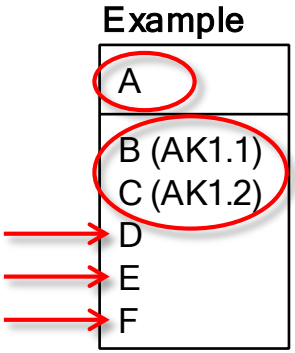
Attr 1
Attr 2 Attr 3 Attr 4 (AK2.1) Attr 5 Attr 6 (AK1.1)

No non-key attributes

Attr 1 Attr 2 Attr 3 (AK2.1)
Attr 4 (AK2.2) Attr 5 (AK2.3) Attr 6 (AK1.1)

Second Normal Form

- **Requirements for Second Normal Form (2NF):**
  - Table must be in First Normal Form
  - Non-key attributes must not depend on subset of any candidate key
    - "Non-key attributes" – not part of *any* candidate key
    - "Subset of candidate key" – requires at least one composite candidate key



$A \rightarrow B$ $A \rightarrow C$ $A \rightarrow D$ $A \rightarrow E$ $A \rightarrow F$

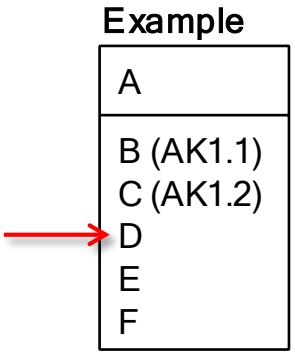
$B \rightarrow E$

$D \rightarrow F$

$\{B, C\} \rightarrow A$ $\{B, C\} \rightarrow D$ $\{B, C\} \rightarrow F$

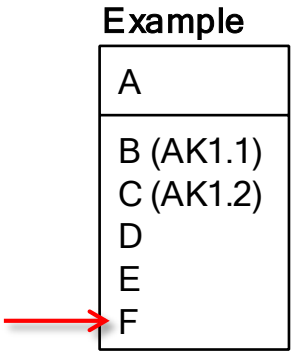
Second Normal Form

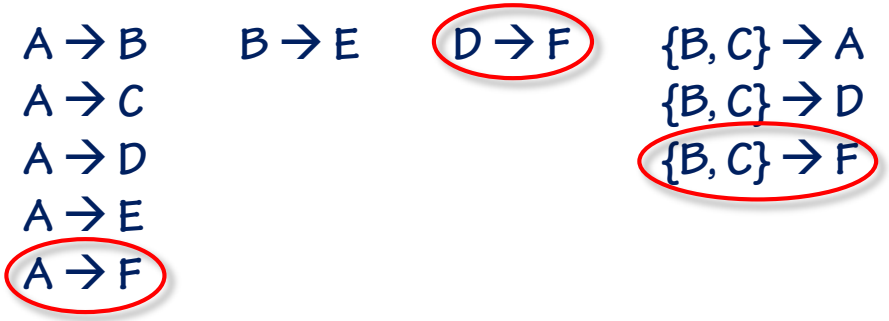
- **Requirements for Second Normal Form (2NF):**
  - Table must be in First Normal Form
  - Non-key attributes must not depend on subset of any candidate key
    - "Non-key attributes" – not part of *any* candidate key
    - "Subset of candidate key" – requires at least one composite candidate key



Second Normal Form

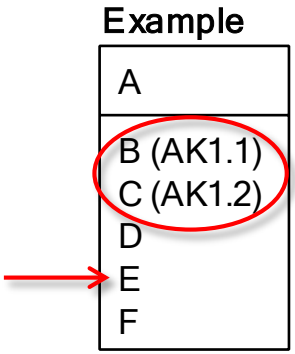
- **Requirements for Second Normal Form (2NF):**
  - Table must be in First Normal Form
  - Non-key attributes must not depend on subset of any candidate key
    - "Non-key attributes" – not part of *any* candidate key
    - "Subset of candidate key" – requires at least one composite candidate key

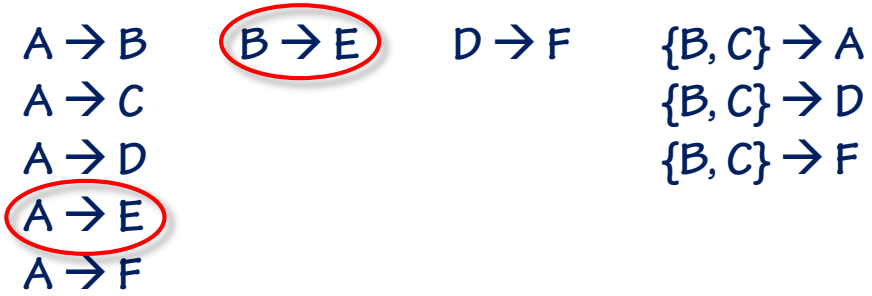




Second Normal Form

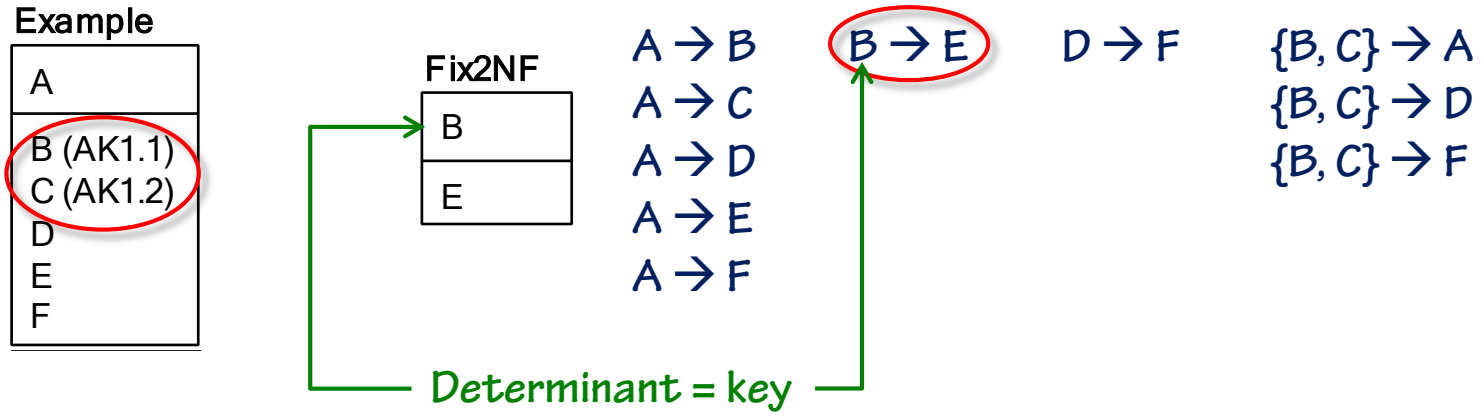
- **Requirements for Second Normal Form (2NF):**
  - Table must be in First Normal Form
  - Non-key attributes must not depend on subset of any candidate key
    - "Non-key attributes" – not part of *any* candidate key
    - "Subset of candidate key" – requires at least one composite candidate key





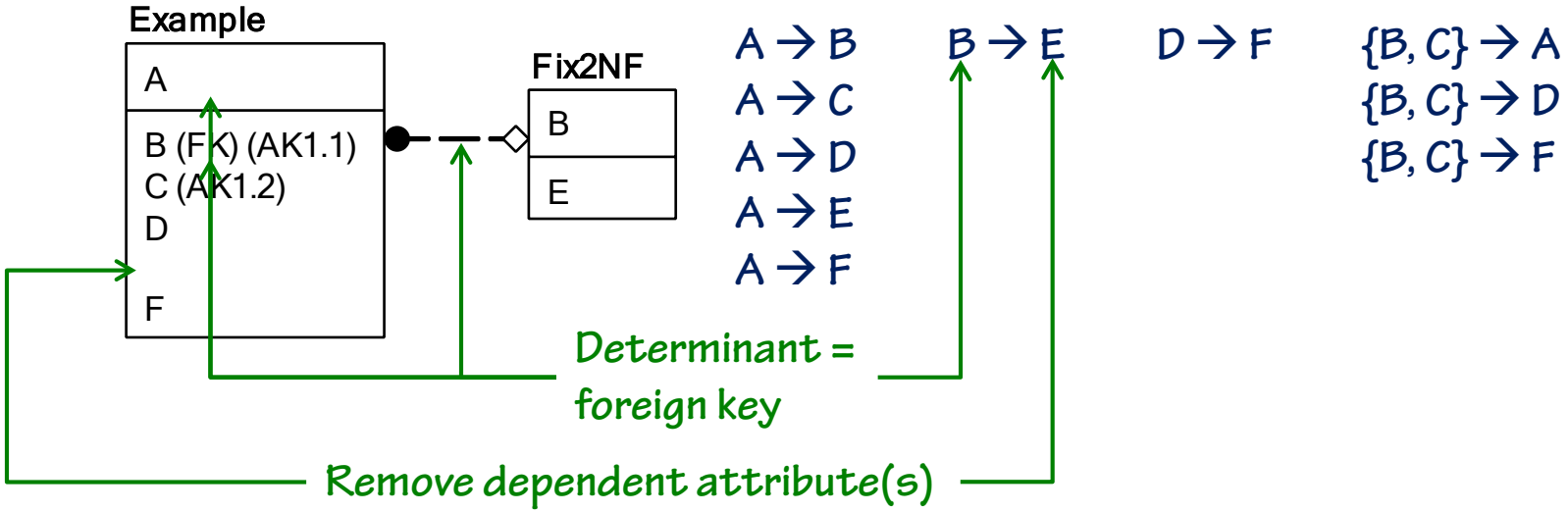
Second Normal Form

- **Requirements for Second Normal Form (2NF):**
  - Table must be in First Normal Form
  - Non-key attributes must not depend on subset of any candidate key
- **Fixing 2NF violations:**
  - New entity type for offending functional dependency



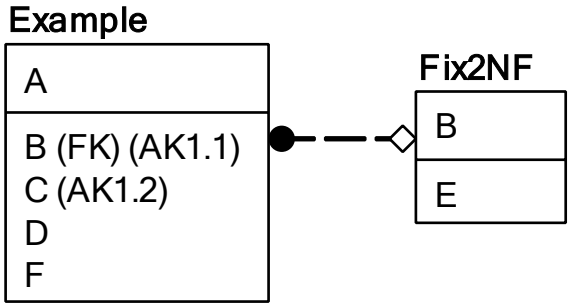
Second Normal Form

- **Requirements for Second Normal Form (2NF):**
  - Table must be in First Normal Form
  - Non-key attributes must not depend on subset of any candidate key
- **Fixing 2NF violations:**
  - New entity type for offending functional dependency
  - Original entity type:
    - Remove dependent attributes
    - Determinant implements relationship



Second Normal Form

- **Requirements for Second Normal Form (2NF):**
  - Table must be in First Normal Form
  - Non-key attributes must not depend on subset of any candidate key
- **Fixing 2NF violations:**
  - New entity type for offending functional dependency
  - Original entity type:
    - Remove dependent attributes
    - Determinant implements relationship



$A \rightarrow B$ $A \rightarrow C$ $A \rightarrow D$ $A \rightarrow E$ $A \rightarrow F$

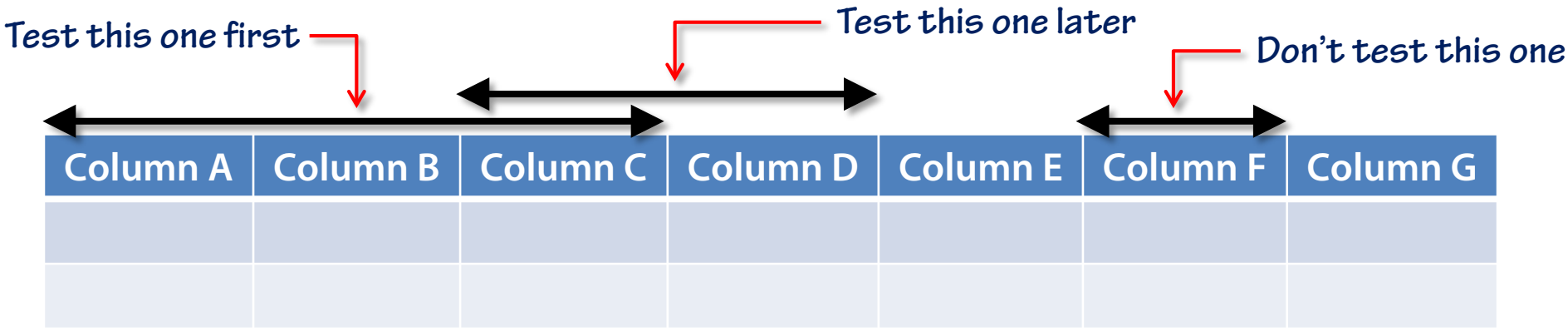
$B \rightarrow E$

$D \rightarrow F$

$\{B, C\} \rightarrow A$ $\{B, C\} \rightarrow D$ $\{B, C\} \rightarrow F$

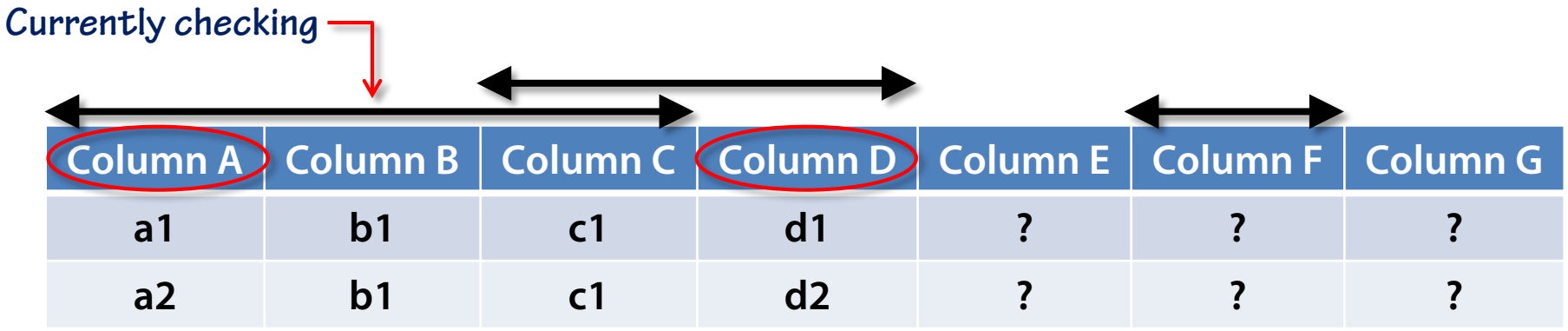
Finding functional dependencies, part 3

- **Dependencies on subsets of candidate keys**
  - Violations of 2NF
    - (and a few extra functional dependencies)
  - Skip for tables with:
    - Only single-column candidate keys
    - One candidate key on *all* columns
      - (many-to-many relationships!)
    - Exactly one composite key **AND** no non-key columns
  - Remaining tables:
    - Procedure executed for each *composite* (multi-column) key



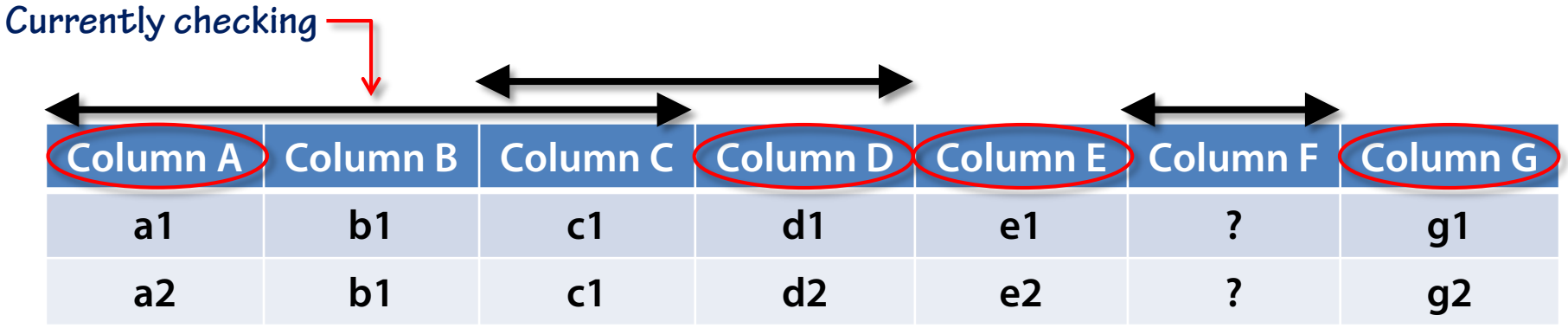
Finding functional dependencies, part 3

- **Dependencies on subsets of candidate keys**
  - Pattern to be populated:
    - Columns in "current" candidate key:
      - One column different, rest equal
    - Columns not in "current" candidate key and not single-column key:
      - One column different, rest irrelevant



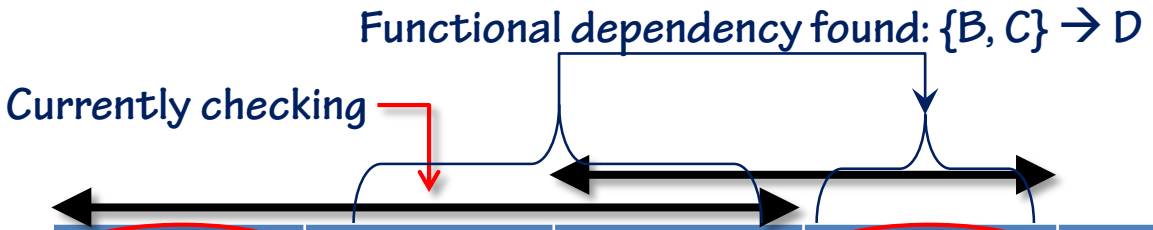
Finding functional dependencies, part 3

- **Dependencies on subsets of candidate keys**
  - Pattern to be populated:
    - Columns in "current" candidate key:
      - One column different, rest equal
    - Columns not in "current" candidate key and not single-column key:
      - One column different, rest irrelevant
      - Optionally combine different columns
        - Faster if there's no dependency, more work if there is



Finding functional dependencies, part 3

- **Dependencies on subsets of candidate keys**
  - Impossible to create **valid** example with the required pattern?
    - Column with difference depends on column(s) with no difference
      - Can be a normal dependency or a derivation rule
    - Don't change data model (yet)



Column A	Column B	Column C	Column D	Column E	Column F	Column G
a1	b1	c1	d1	?	?	?
a2	b1	c1	d2	?	?	?

Finding functional dependencies, part 3

- **Dependencies on subsets of candidate keys**
  - Impossible to create **valid** example with the required pattern?
    - Column with difference depends on column(s) with no difference
      - Can be a normal dependency or a derivation rule
    - Don't change data model (yet)
    - Continue testing
      - Other columns

Functional dependency found: {B, C} → D

Currently checking

Column A	Column B	Column C	Column D	Column E	Column F	Column G
a1	b1	c1	?	?	?	g1
a2	b1	c1	?	?	?	g2

Finding functional dependencies, part 3

- **Dependencies on subsets of candidate keys**
  - Impossible to create **valid** example with the required pattern?
    - Column with difference depends on column(s) with no difference
      - Can be a normal dependency or a derivation rule
    - Don't change data model (yet)
    - Continue testing
      - Other columns
      - Other subsets of current candidate key

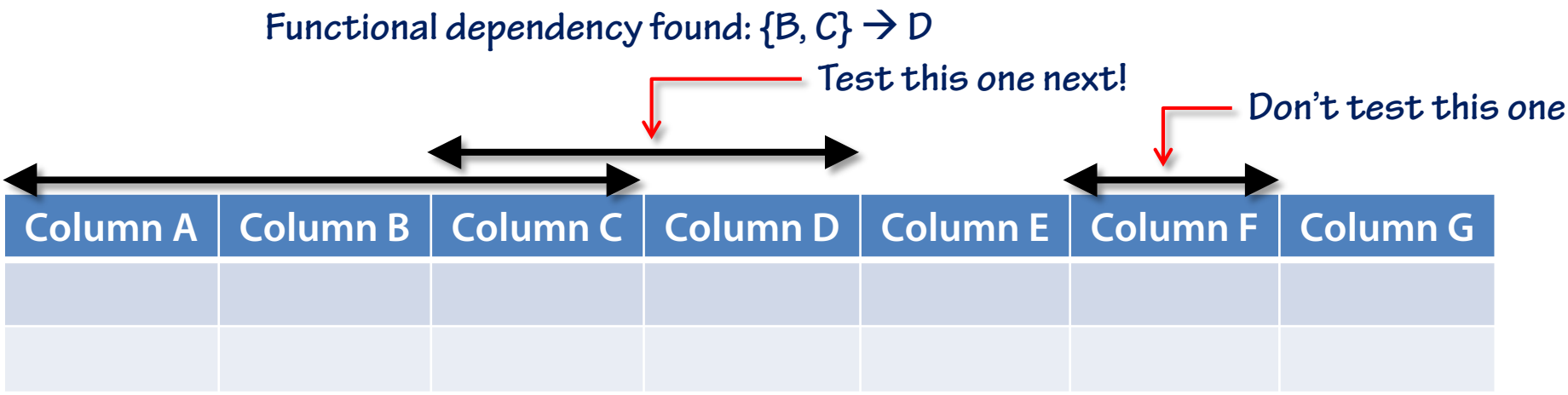
Functional dependency found: {B, C} → D

Currently checking

Column A	Column B	Column C	Column D	Column E	Column F	Column G
a1	b1	c1				
a1	b1	c2				

Finding functional dependencies, part 3

- **Dependencies on subsets of candidate keys**
  - Impossible to create **valid** example with the required pattern?
    - Column with difference depends on column(s) with no difference
      - Can be a normal dependency or a derivation rule
    - Don't change data model (yet)
    - Continue testing
      - Other columns
      - Other subsets of current candidate key
      - Other composite candidate keys



Finding functional dependencies, part 3

- **Dependencies on subsets of candidate keys**
  - All functional dependencies on subset of keys found
    - Might not be full!

Actual dependency: A → E

Dependencies found: {A, B, C} → E
{A, B, D} → E
{A, C, D} → E

Actual dependency: {A, B} → E

Dependencies found: {A, B, C} →
{A, B, D} → E

Column A	Column B	Column C	Column D	Column E	Column F	Column G

Finding functional dependencies, part 3

- **Dependencies on subsets of candidate keys**
  - All functional dependencies on subset of keys found
    - Might not be full!
    - Look for common subset …
      - … for which every superset was found

Actual dependency: A → E

Actual dependencies: {A, B} →
{A, C, D} → E

Actual dependencies: {A, C} →
{A, B, D} → E

Actual dependencies: {A, D} →
{A, B, C} → E

Dependencies found: {A, B, C} → E
{A, B, D} → E
{A, C, D} → E

Column A	Column B	Column C	Column D	Column E	Column F	Column G

Finding functional dependencies, part 3

- **Dependencies on subsets of candidate keys**
  - All functional dependencies on subset of keys found
    - Might not be full!
    - Look for common subset …
      - … for which every superset was found
      - Multiple common subsets possible

Actual dependencies: {A, B} → E
{A, C} → E

Actual dependencies: {A, B} → E
{A, D} → E

Actual dependencies: {A, C} → E
{A, D} → E

Actual dependencies: {A, B} → E
{A, C} → E
{A, D} → E

Dependencies found: {A, B, C} → E
{A, B, D} → E
{A, C, D} → E

Column A	Column B	Column C	Column D	Column E	Column F	Column G

Finding functional dependencies, part 3

- **Dependencies on subsets of candidate keys**
  - All functional dependencies on subset of keys found
    - Might not be full!
    - Look for common subset …
      - … for which **every** superset was found
      - Multiple common subsets possible

Not a dependency: {B, C} → E
would have implied: {B,C, D} → E

Dependencies found: {A, B, C} → E
{A, B, D} → E
{A, C, D} → E

Not a dependency: B → E
would have implied: {B, C, D} → E
{B, C} → E
etcetera

Column A	Column B	Column C	Column D	Column E	Column F	Column G

Finding functional dependencies, part 3

- **Dependencies on subsets of candidate keys**
  - All functional dependencies on subset of keys found
    - Might not be full!
    - Look for common subset …
      - … for which **every** superset was found
      - Multiple common subsets possible
      - Confirmation (by testing) required
        - If confirmed: Add new dependency, remove implied dependencies
      - Check all possibilities!

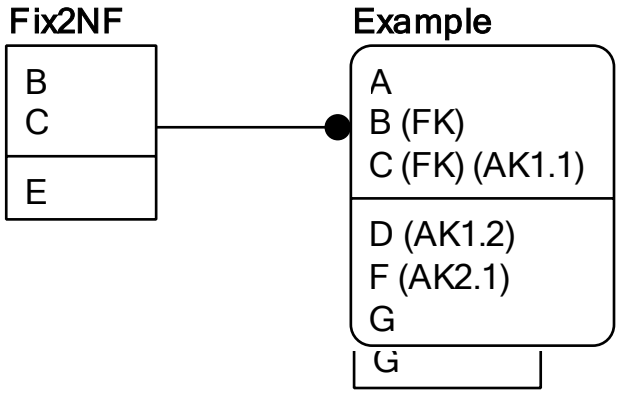
Dependencies found: {A, B, C} → E
{A, B, D} → E
{A, C, D} → E

Possible dependencies: A → E
{A, B} → E
{A, C} → E
{A, D} → E

Column A	Column B	Column C	Column D	Column E	Column F	Column G
a1	b1	c1	d1	e1	?	?
a1	b2	c1	d2	e2	?	?

Finding functional dependencies, part 3

- **Dependencies on subsets of candidate keys**
  - All ***full*** functional dependencies on subset of keys found
    - Caused by derivation rule? → handled later
    - Non-key column depends on subset of key → violates 2NF
      - Create new entity type
      - Remove dependent attribute



Full functional dependencies found: {B, C} → E

D → A

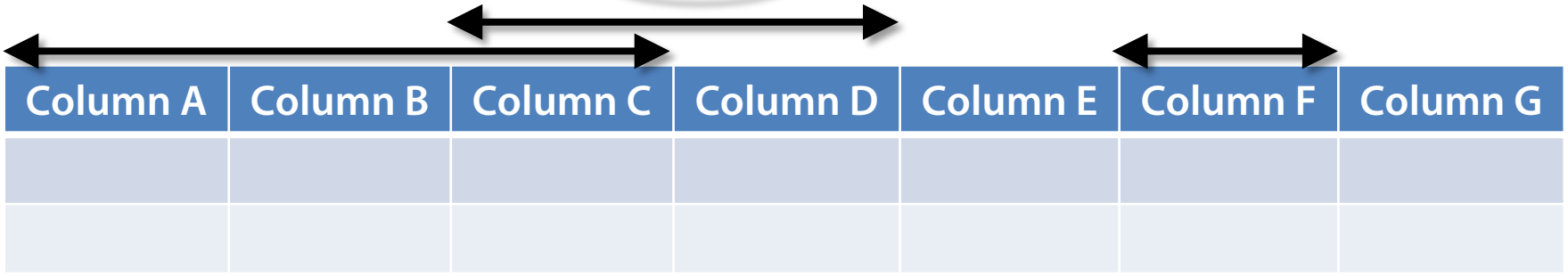
Column A	Column B	Column C	Column D	Column E	Column F	Column G

Finding functional dependencies, part 3

- **Dependencies on subsets of candidate keys**
  - All ***full*** functional dependencies on subset of keys found
    - Caused by derivation rule? → handled later
    - Non-key column depends on subset of key → violates 2NF
      - Create new entity type
      - Remove dependent attribute
    - Subset of one key depends on subset of another key → No 2NF violation
      - No schema change required

Full functional dependencies found: {B, C} → E

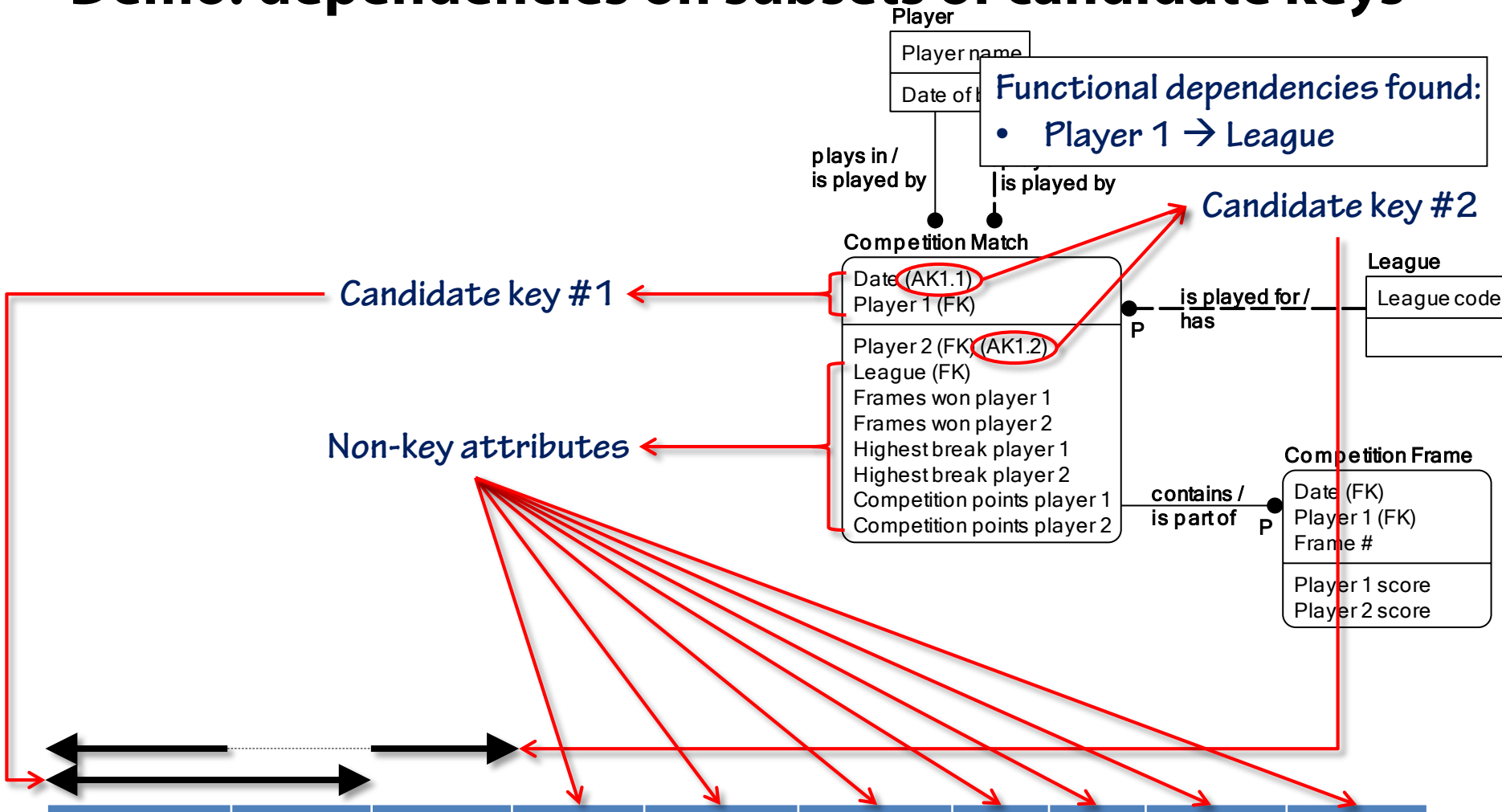
D → A



Finding functional dependencies, part 3

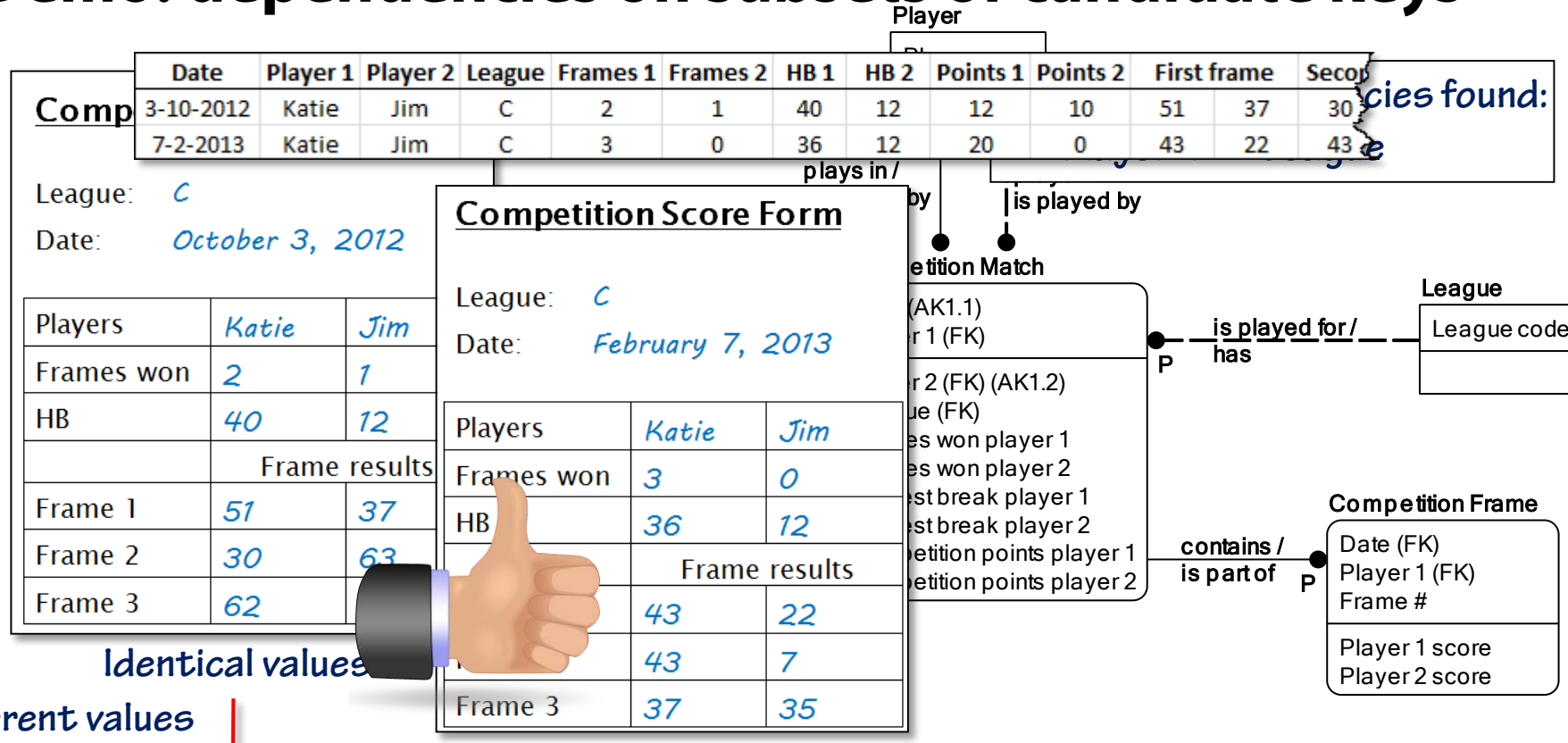
- **Dependencies on subsets of candidate keys**
  - All ***full*** functional dependencies on subset of keys found
    - Caused by derivation rule? → handled later
    - Non-key column depends on subset of key → violates 2NF
      - Create new entity type
      - Remove dependent attribute
    - Subset of one key depends on subset of another key → No 2NF violation
      - No schema change required
    - ***(Rare)*** One non-key column depends on two (or more) subset of keys
      - New entity type for each dependency
        - Causes redundancy
        - Allows data that violates dependencies
      - New entity type for one of the dependencies
        - Allows data that violates dependencies
      - Leave original entity type unchanged
        - Violates 2NF

Demo: dependencies on subsets of candidate keys



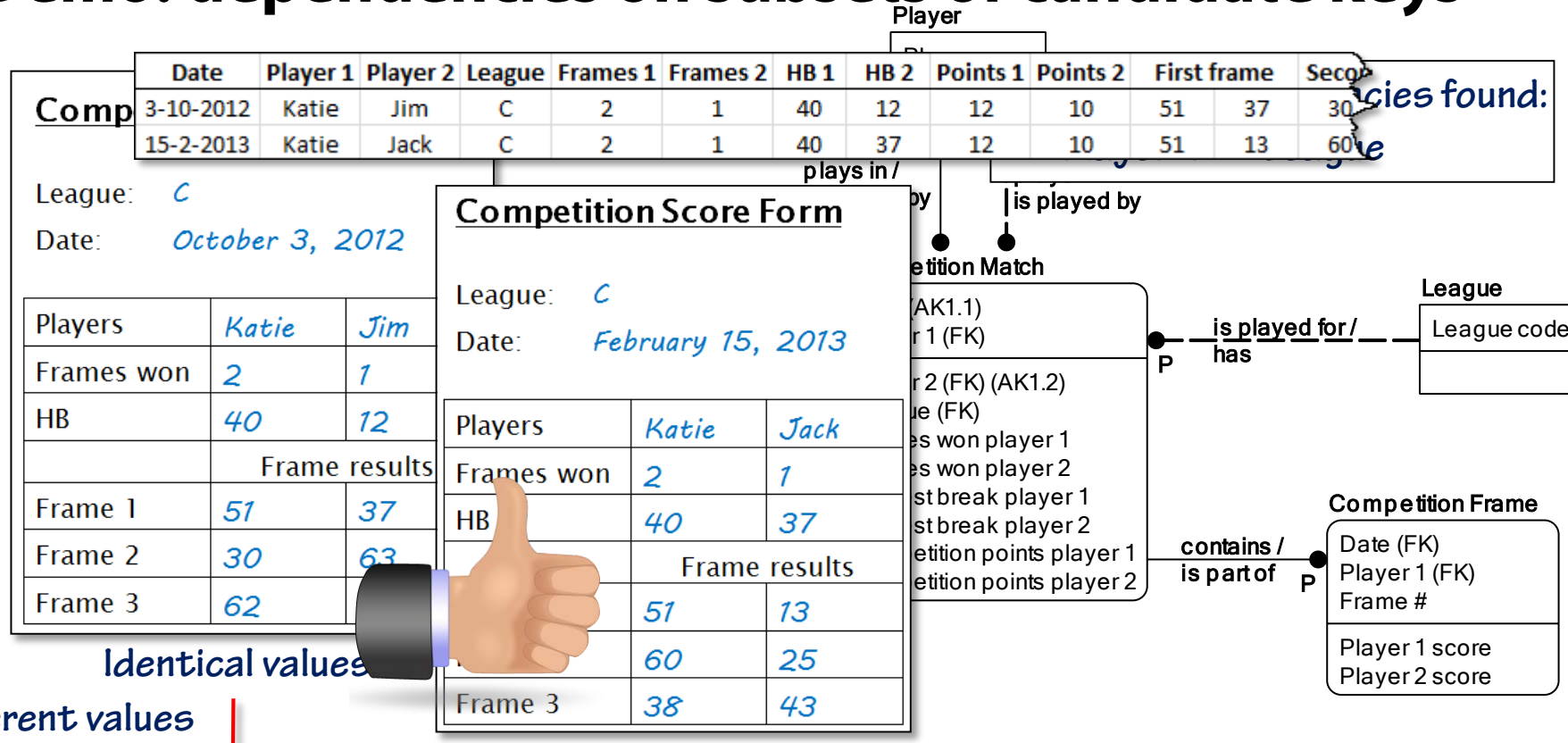
Date	Player 1	Player 2	League	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2

Demo: dependencies on subsets of candidate keys

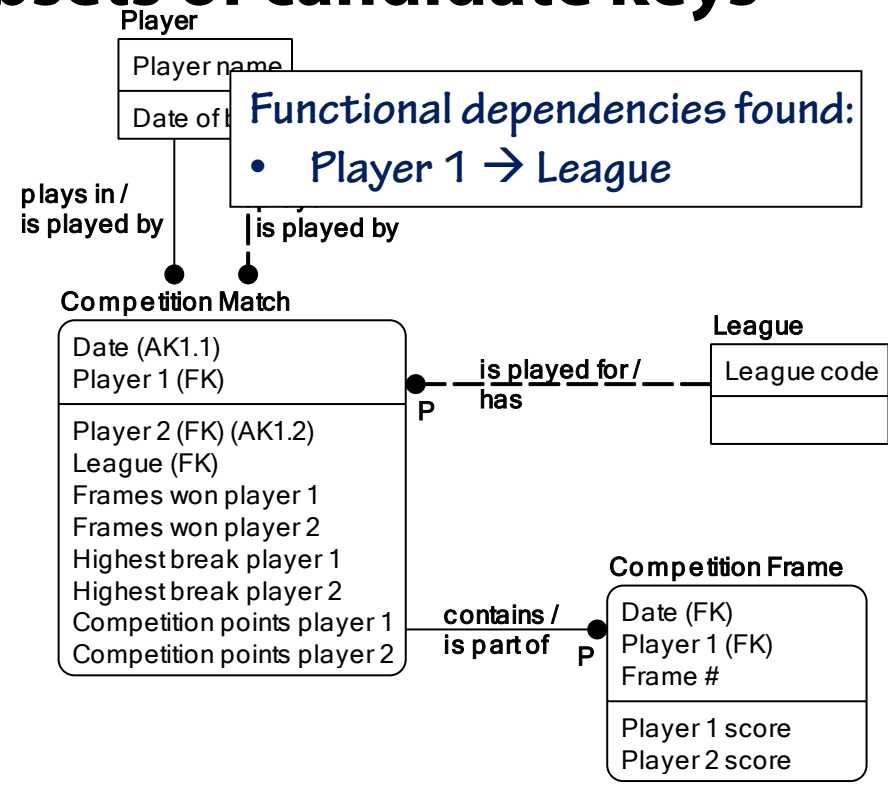


Date	Player 1	Player 2	League	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2
2012-10-03	Katie	Jim	C	2	1	40	12	12	10
2013-02-07	Katie	Jim	C	3	0	36	12	20	0

Demo: dependencies on subsets of candidate keys



Demo: dependencies on subsets of candidate keys



Identical values

Different values

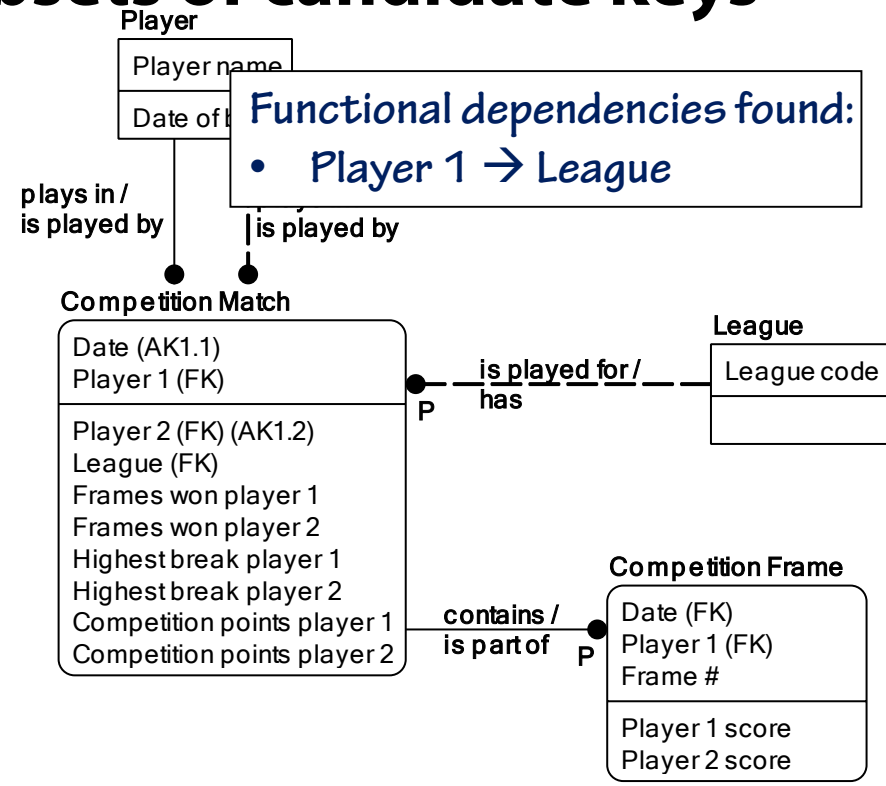
Currently checking

Different values

?

Date	Player 1	Player 2	League	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2
2012-10-03	Katie	Jim	C	2	1	40	12	12	10
2013-02-15	Katie	Jim	C	2	1	40	37	12	10

Demo: dependencies on subsets of candidate keys



Identical values

Different values

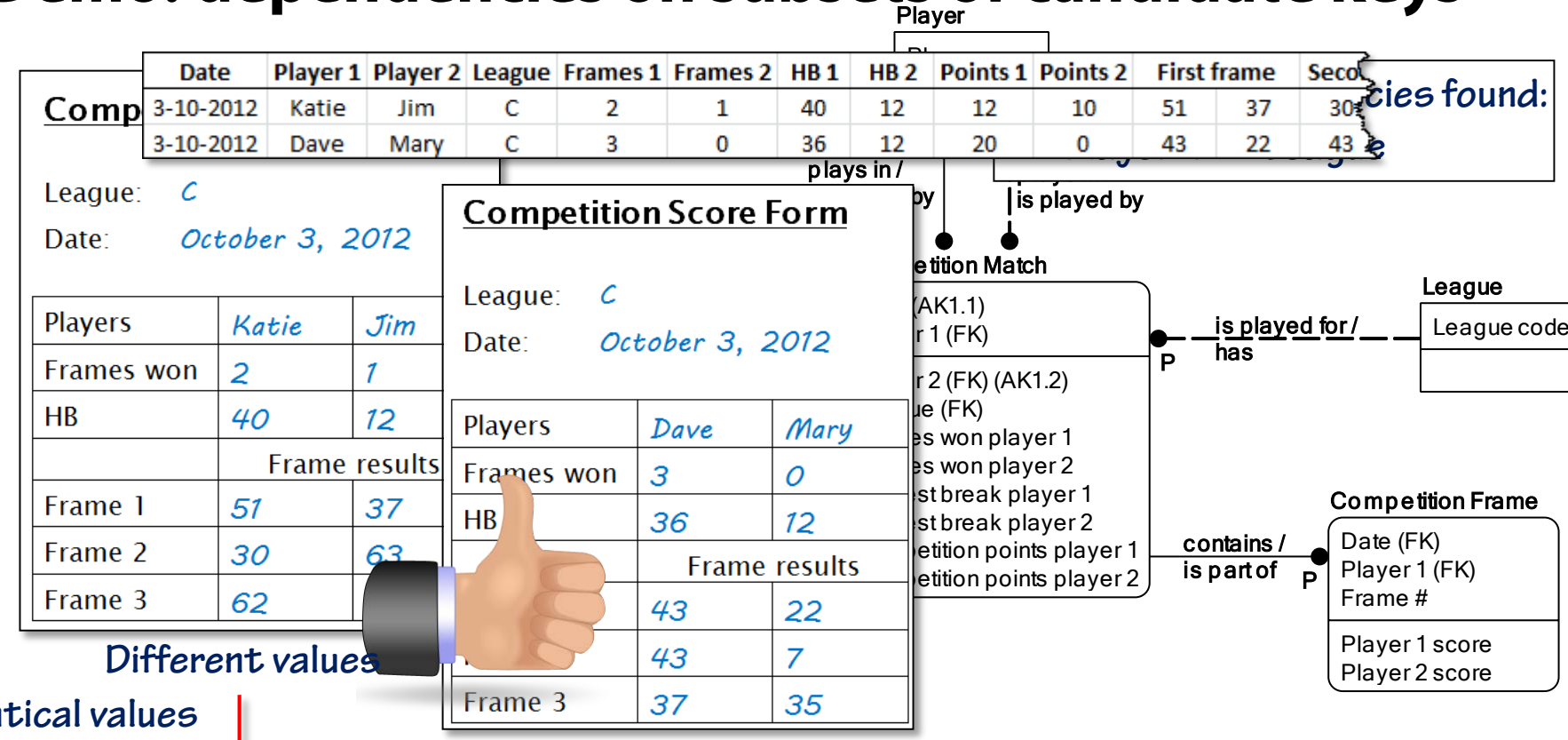
Currently checking

Different values

?

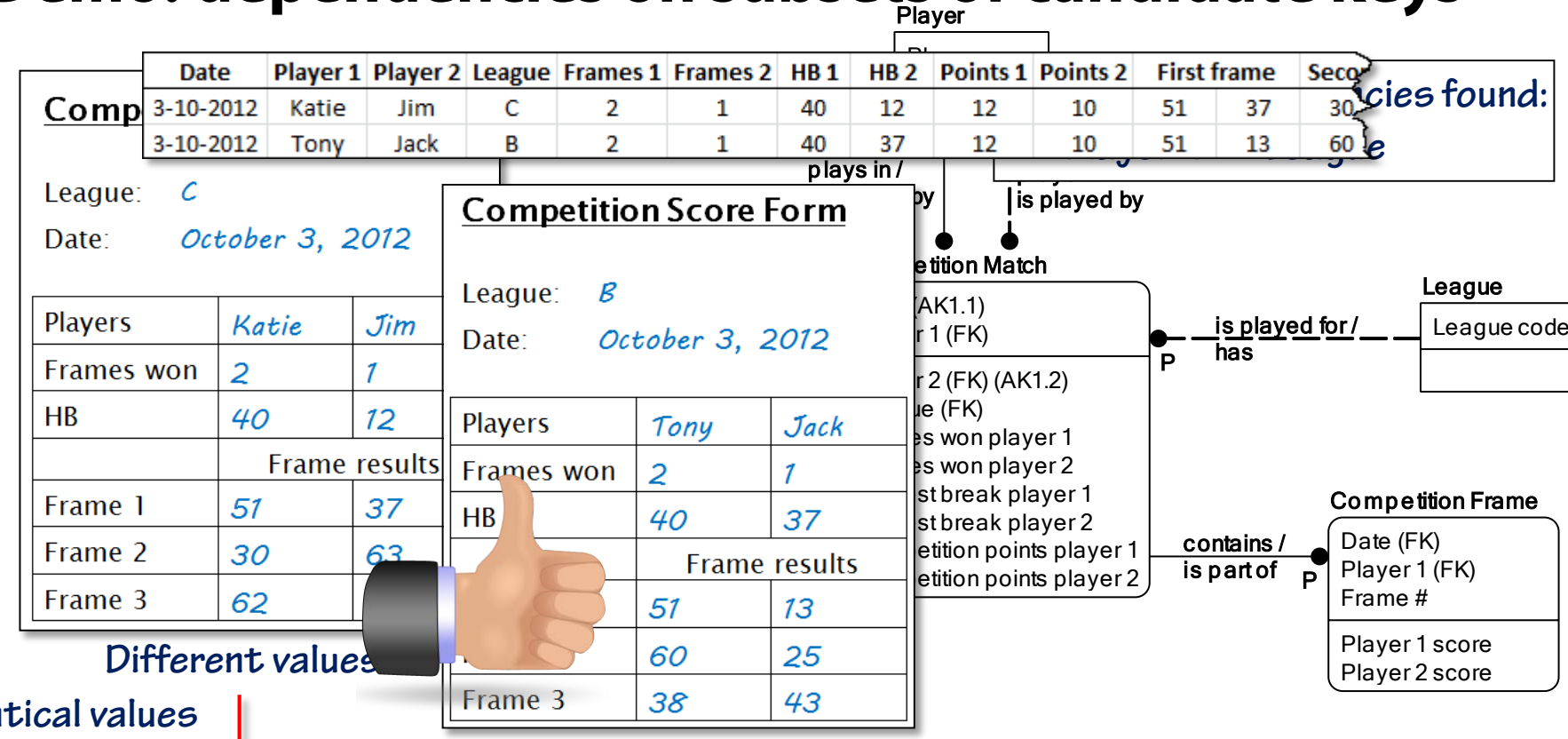
Date	Player 1	Player 2	League	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2
2012-10-03	Katie	Jim	C	2	1	40	12	12	10
2013-02-15	Katie	Jim	C	2	1	40	12	12	10

Demo: dependencies on subsets of candidate keys



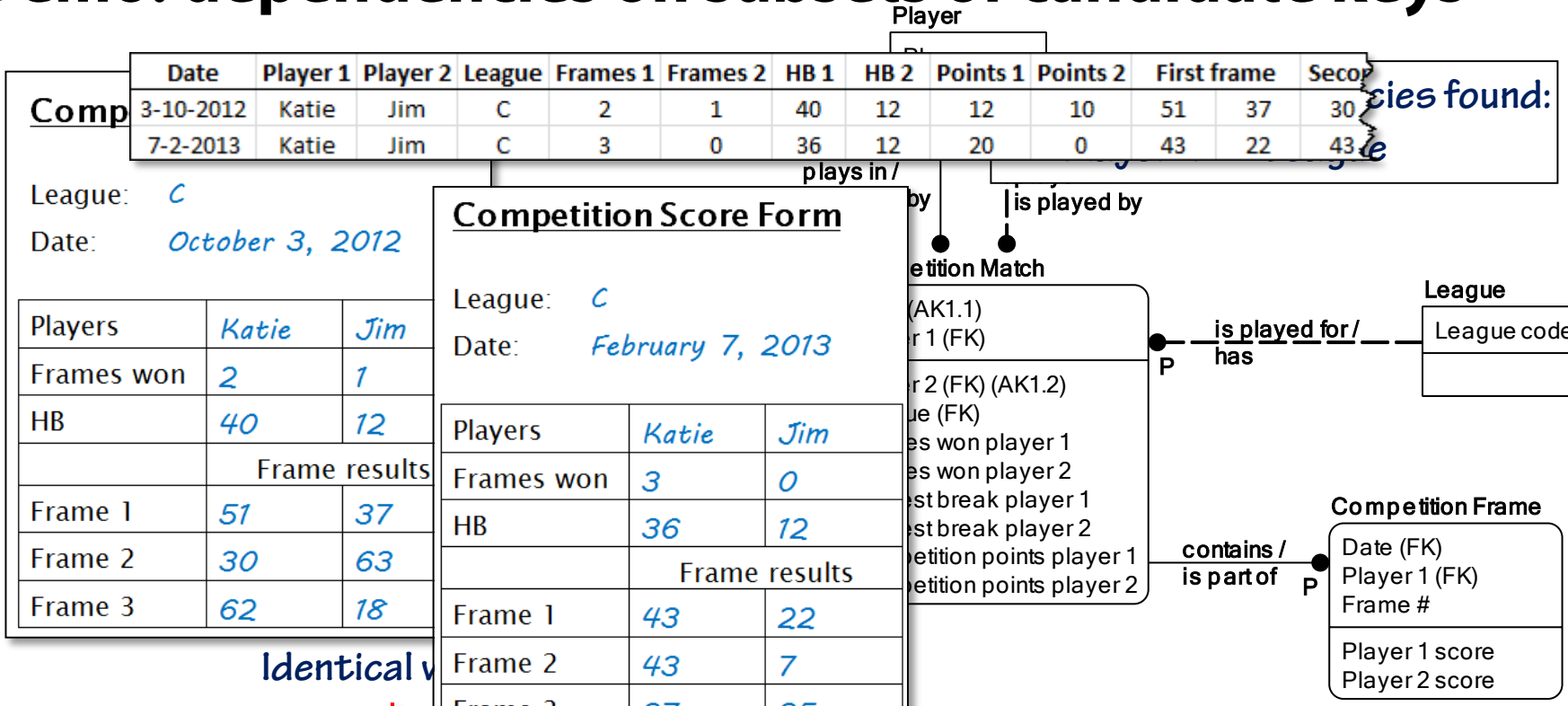
Date	Player 1	Player 2	League	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2
2012-10-03	Katie	Jim	C	2	1	40	12	12	10
2012-10-03	Dave	Mary	C	3	0	36	12	20	0

Demo: dependencies on subsets of candidate keys



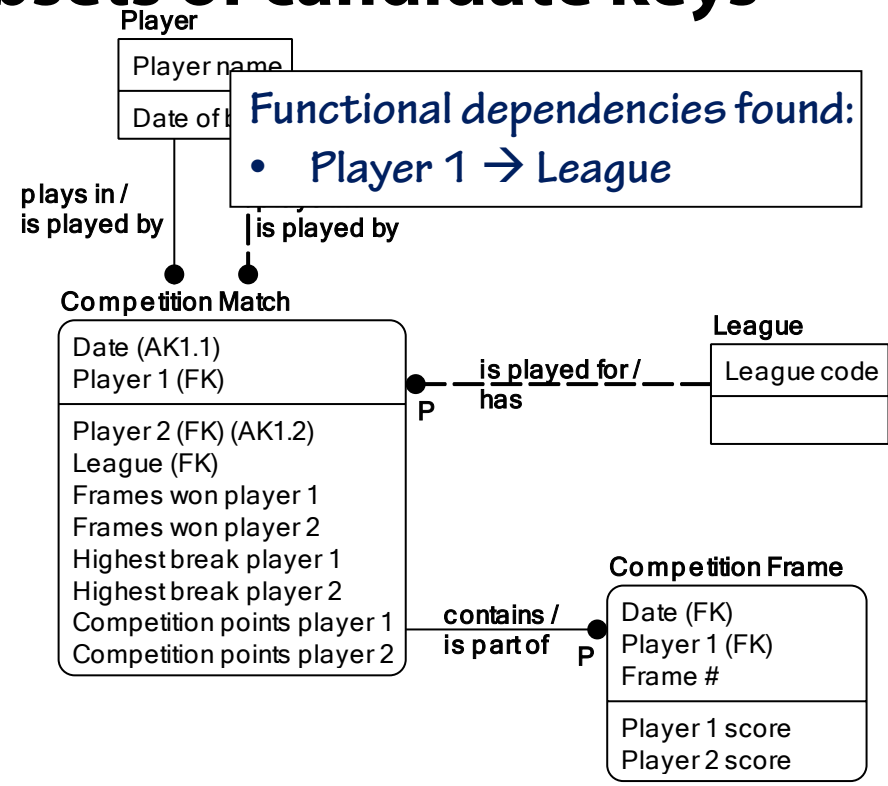
Date	Player 1	Player 2	League	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2
2012-10-03	Katie	Jim	C	2	1	40	12	12	10
2012-10-03	Tony	Jack	B	2	1	40	37	12	10

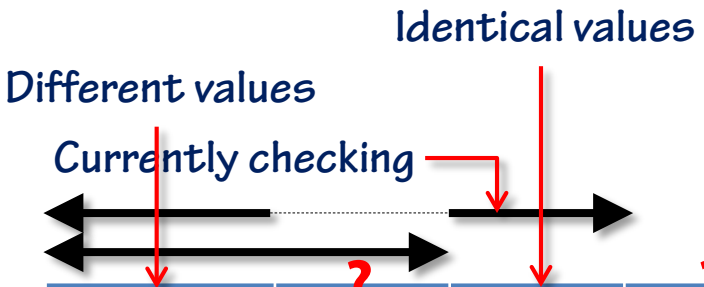
Demo: dependencies on subsets of candidate keys



Date	Player 1	Player 2	League	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2
2012-10-03	Katie	Jim	C	2	1	40	12	12	10
2013-02-07	Katie	Jim	C	3	0	36	12	20	0

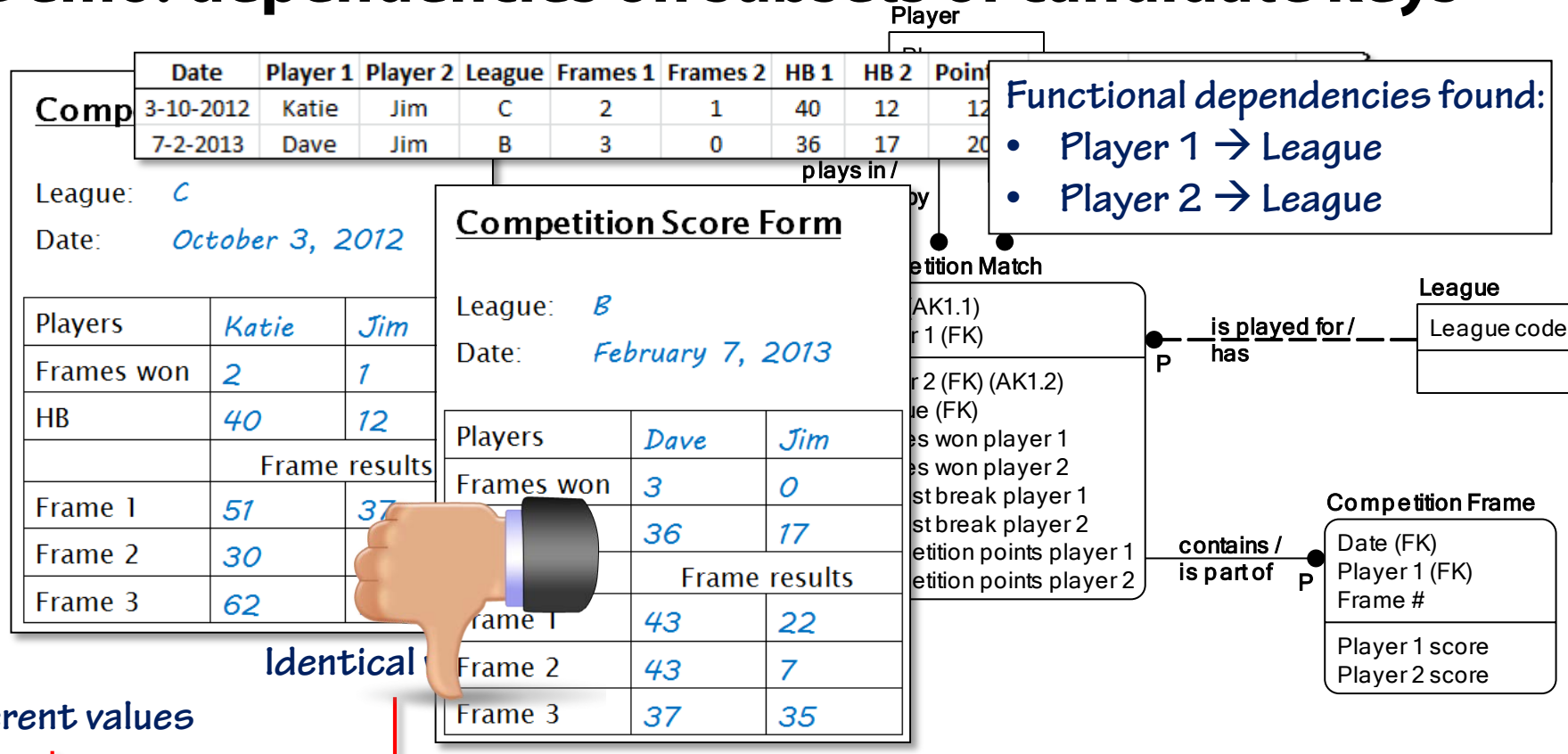
Demo: dependencies on subsets of candidate keys





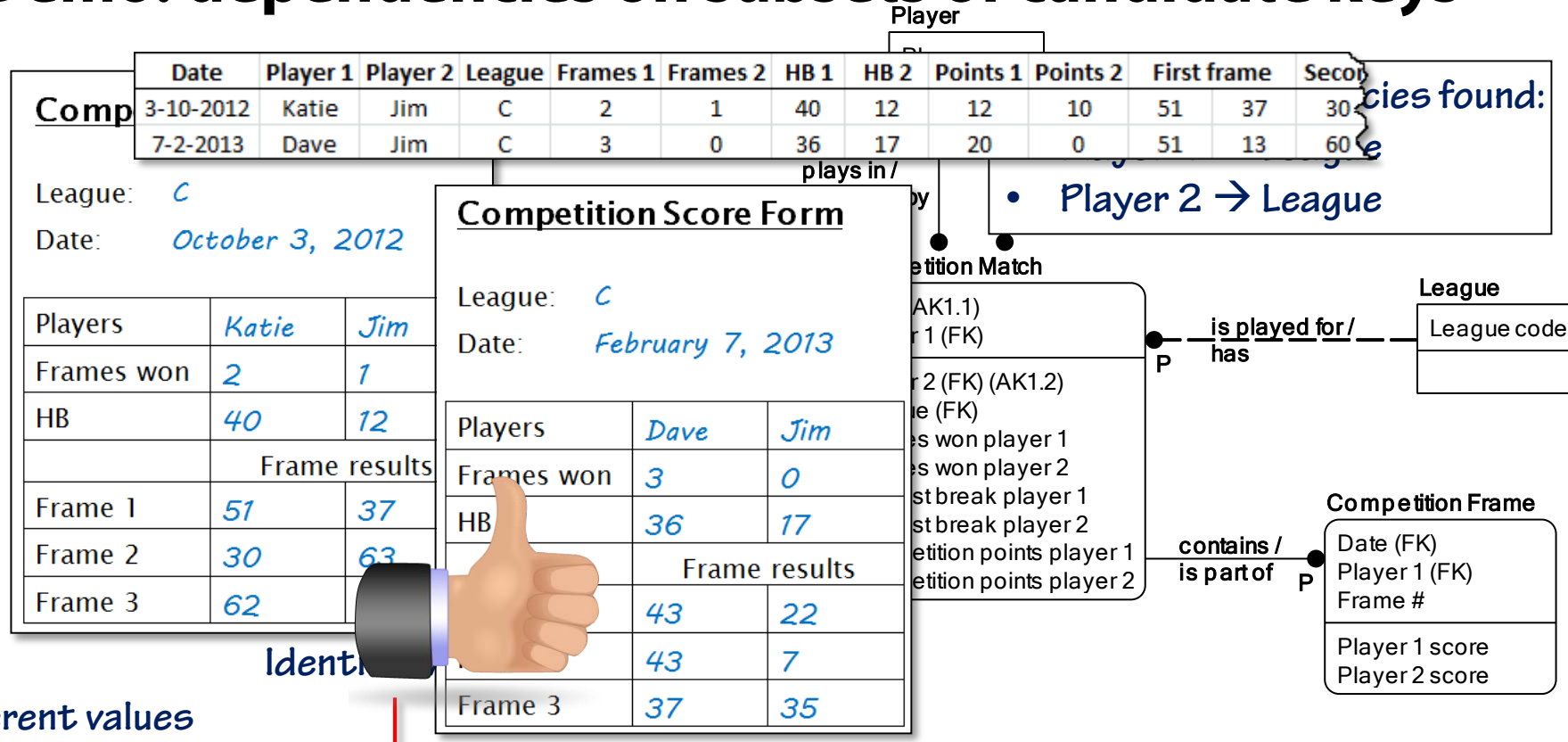
Date	Player 1	Player 2	League	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2
2012-10-03	?	Jim	?	?	?	?	12	?	?
2013-02-07	?	Jim	?	?	?	?	17	?	?

Demo: dependencies on subsets of candidate keys



Date	Player 1	Player 2	League	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2
2012-10-03	Katie	Jim	C	2	1	40	12	12	10
2013-02-07	Dave	Jim	B	3	0	36	17	20	0

Demo: dependencies on subsets of candidate keys



Date	Player 1	Player 2	League	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2
2012-10-03	Katie	Jim	C	2	1	40	12	12	10
2013-02-07	Dave	Jim	C	3	0	36	17	20	0

Demo: dependencies on subsets of candidate keys

Date	Player 1	Player 2	League	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2	First frame	Seco
3-10-2012	Katie	Jim	C	2	1	40	12	12	10	51	37	30
3-10-2012	Dave	Mary	C	3	0	36	12	20	0	43	22	43

cies found:

- Player 2 → League

Comp

League: C

Date: October 3, 2012

Players	Katie	Jim
Frames won	2	1
HB	40	12
	Frame results	
Frame 1	51	37
Frame 2	30	63
Frame 3	62	

Differ

Competition Score Form

League: C

Date: October 3, 2012

Players	Dave	Mary
Frames won	3	0
HB	36	12
	Frame results	
	43	22
	43	7
Frame 3	37	35

Player

plays in /

by

etition Match

(AK1.1)
r 1 (FK)

r 2 (FK) (AK1.2)
ue (FK)
es won player 1
es won player 2
st break player 1
st break player 2
etition points player 1
etition points player 2

P

is played for / has

League

League code

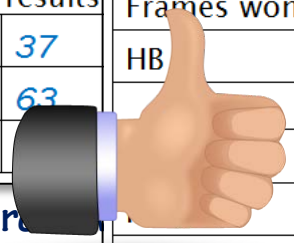
contains / is part of

P

Competition Frame

Date (FK)
Player 1 (FK)
Frame #

Player 1 score
Player 2 score

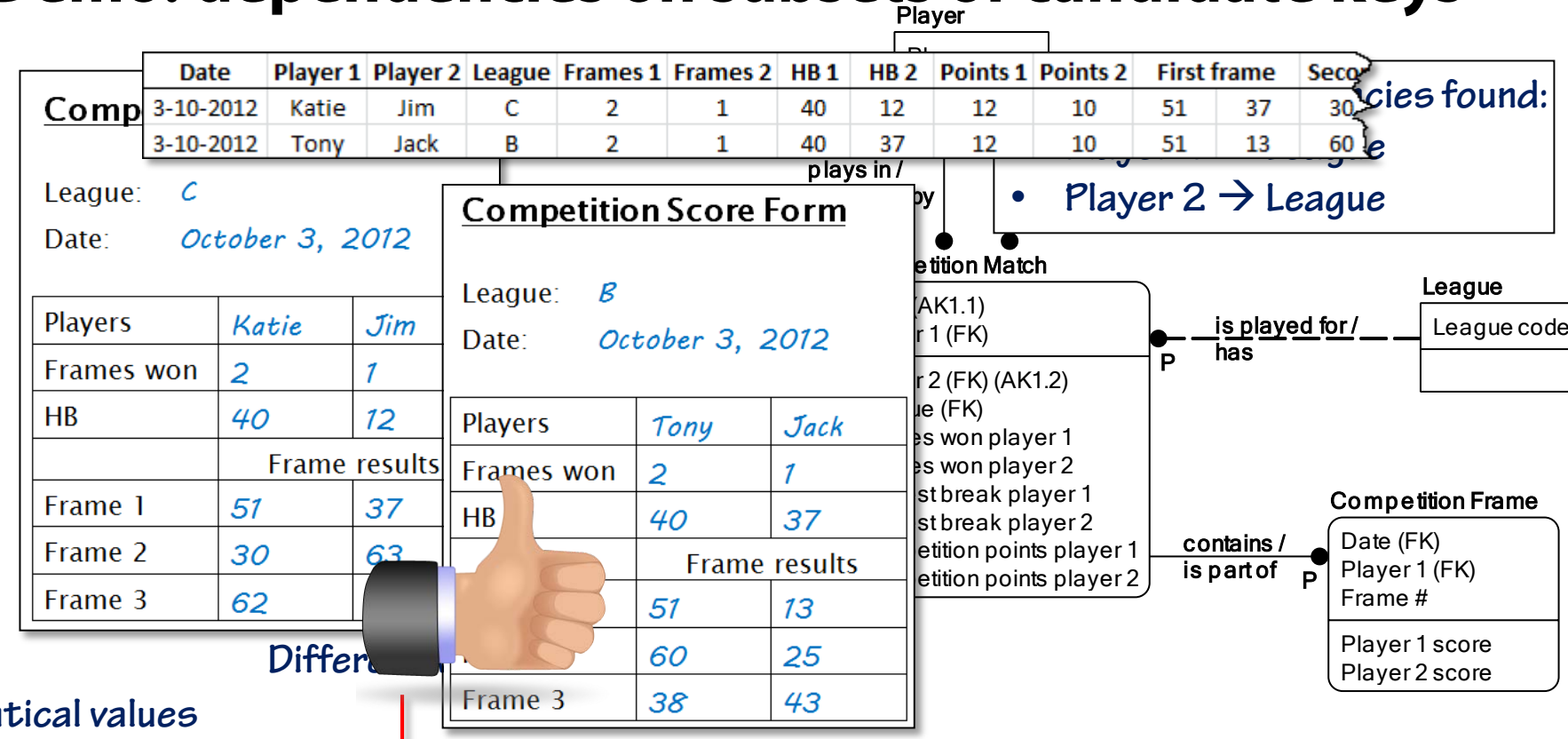


Identical values

Currently checking

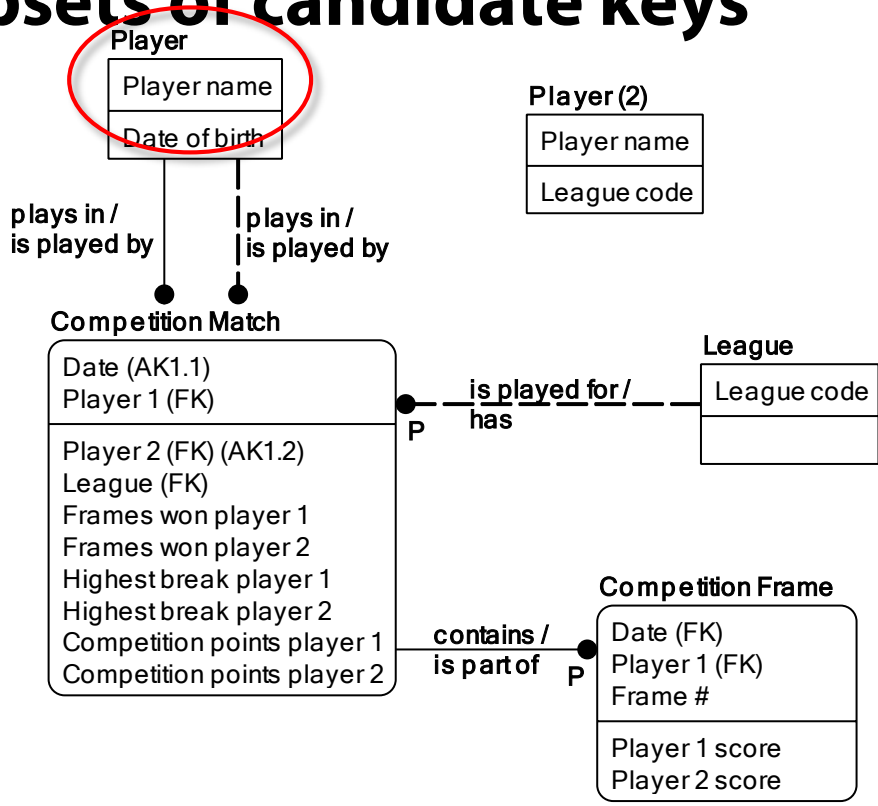
Date	Player 1	Player 2	League	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2
2012-10-03	Katie	Jim	C	2	1	40	12	12	10
2012-10-03	Dave	Mary	C	3	0	36	12	20	0

Demo: dependencies on subsets of candidate keys



Date	Player 1	Player 2	League	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2
2012-10-03	Katie	Jim	C	2	1	40	12	12	10
2012-10-03	Tony	Jack	B	2	1	40	37	12	10

Demo: dependencies on subsets of candidate keys



Functional dependencies found:

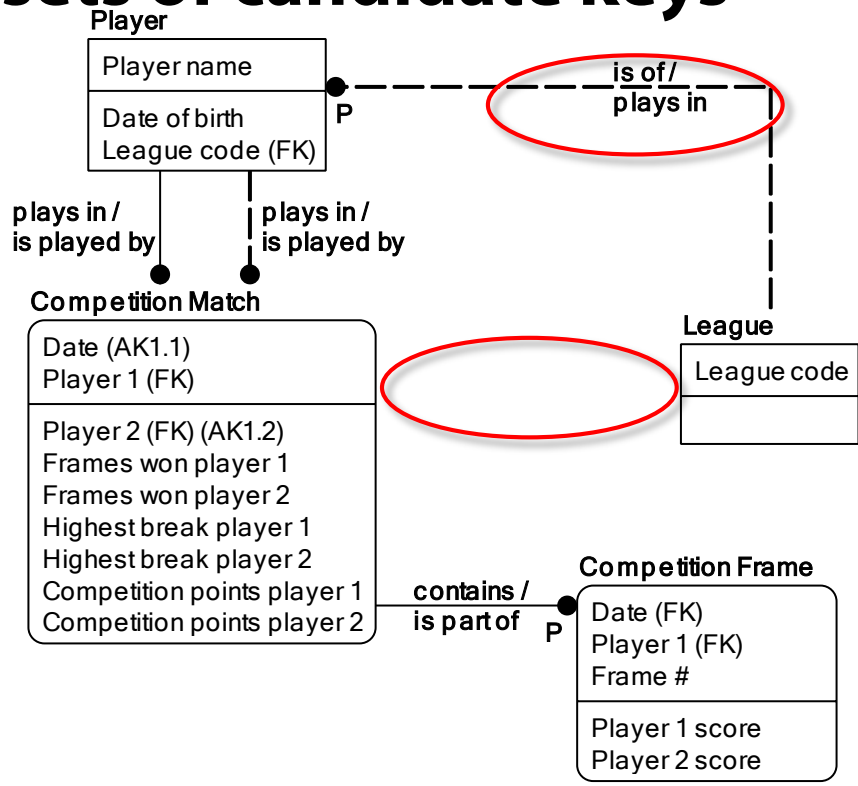
- Player 1 → League
- Player 2 → League

Actual functional dependency:

- Player → League

Date	Player 1	Player 2	League	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2

Demo: dependencies on subsets of candidate keys



Functional dependencies found:

- Player 1 → League
- Player 2 → League

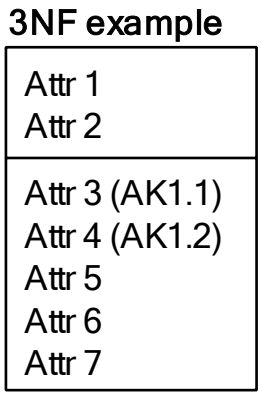
Actual functional dependency:

- Player → League

Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2

Third Normal Form

- **Requirements for Third Normal Form (3NF):**
  - Table must be in Second Normal Form
  - Non-prime attributes must be non-transitively dependent on every superkey
    - or: No dependency of non-key attribute on attribute(s) that is/are not a key
      - A non-key attribute depends on one non-key attribute

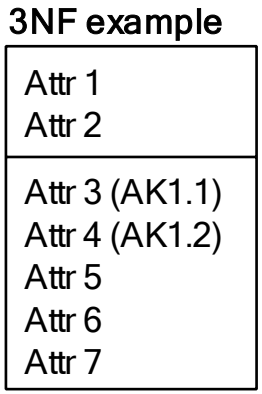


Examples of violations:

- Attr5 → Attr7
- Attr7 → Attr6

Third Normal Form

- **Requirements for Third Normal Form (3NF):**
  - Table must be in Second Normal Form
  - Non-prime attributes must be non-transitively dependent on every superkey
    - or: No dependency of non-key attribute on attribute(s) that is/are not a key
      - A non-key attribute depends on one non-key attribute
      - A non-key attribute depends on multiple non-key attributes

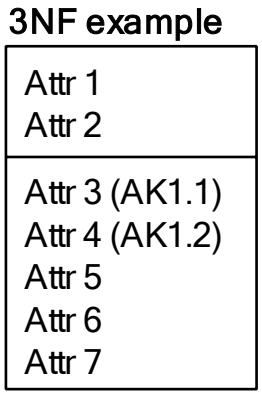


Examples of violations:

- {Attr5, Attr6} → Attr7
- {Attr5, Attr7} → Attr6

Third Normal Form

- **Requirements for Third Normal Form (3NF):**
  - Table must be in Second Normal Form
  - Non-prime attributes must be non-transitively dependent on every superkey
    - or: No dependency of non-key attribute on attribute(s) that is/are not a key
      - A non-key attribute depends on one non-key attribute
      - A non-key attribute depends on multiple non-key attributes
      - A non-key attribute depends on combination of candidate key subset and non-key

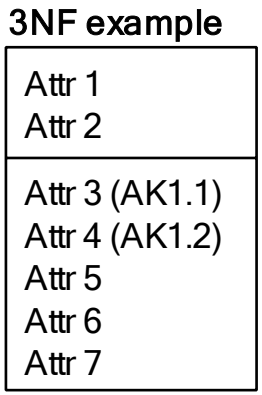


Examples of violations:

- {Attr1, Attr5} → Attr7
- {Attr3, Attr5, Attr7} → Attr6

Third Normal Form

- **Requirements for Third Normal Form (3NF):**
  - Table must be in Second Normal Form
  - Non-prime attributes must be non-transitively dependent on every superkey
    - or: No dependency of non-key attribute on attribute(s) that is/are not a key
      - A non-key attribute depends on one non-key attribute
      - A non-key attribute depends on multiple non-key attributes
      - A non-key attribute depends on combination of candidate key subset and non-key
      - A non-key attribute depends on combination of multiple candidate key subsets

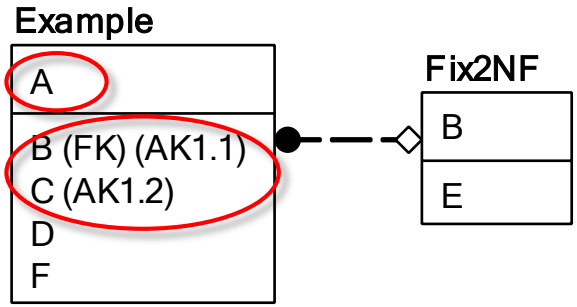


Examples of violations:

- {Attr1, Attr4} → Attr6
- {Attr2, Attr3} → Attr6
- {Attr2, Attr3, Attr5, Attr6} → Attr7

Third Normal Form

- **Requirements for Third Normal Form (3NF):**
  - Table must be in Second Normal Form
  - Non-prime attributes must be non-transitively dependent on every superkey
    - or: No dependency of non-key attribute on attribute(s) that is/are not a key
      - A non-key attribute depends on one non-key attribute
      - A non-key attribute depends on multiple non-key attributes
      - A non-key attribute depends on combination of candidate key subset and non-key
      - A non-key attribute depends on combination of multiple candidate key subsets



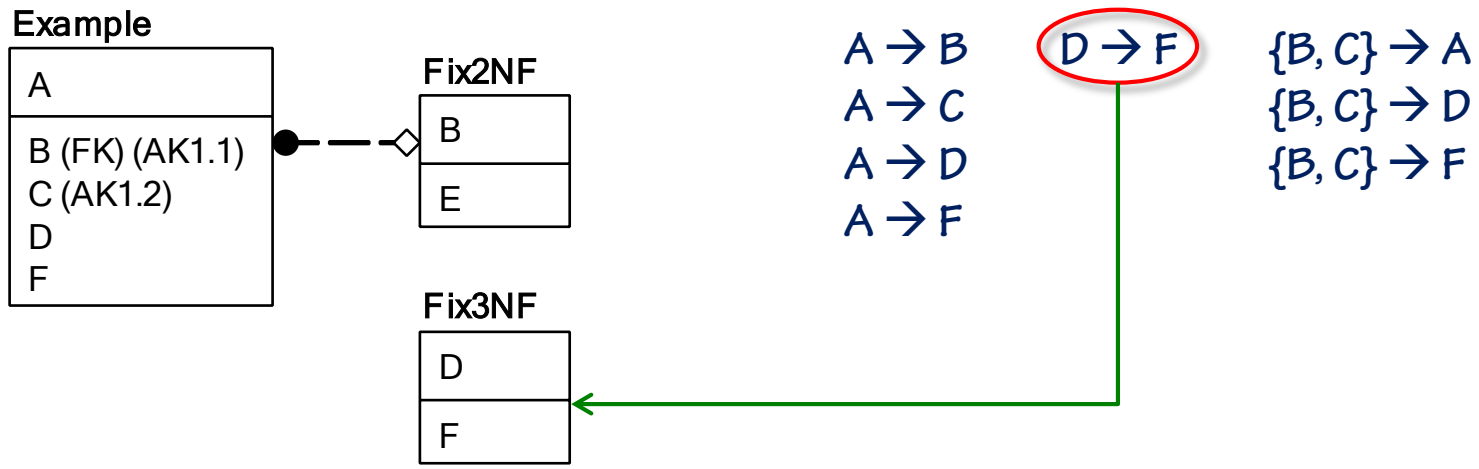
$A \rightarrow B$
$A \rightarrow C$
$A \rightarrow D$
$A \rightarrow F$

$D \rightarrow F$

$\{B, C\} \rightarrow A$
$\{B, C\} \rightarrow D$
$\{B, C\} \rightarrow F$

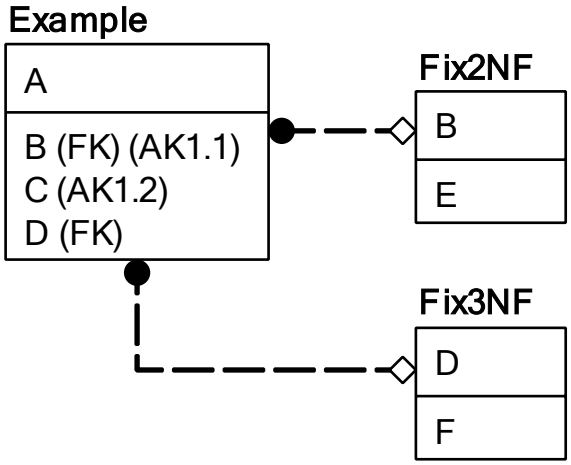
Third Normal Form

- **Requirements for Third Normal Form (3NF):**
  - Table must be in Second Normal Form
  - Non-prime attributes must be non-transitively dependent on every superkey
    - or: No dependency of non-key attribute on attribute(s) that is/are not a key
- **Fixing 3NF violations:**
  - New entity type for offending functional dependency



Third Normal Form

- **Requirements for Third Normal Form (3NF):**
  - Table must be in Second Normal Form
  - Non-prime attributes must be non-transitively dependent on every superkey
    - or: No dependency of non-key attribute on attribute(s) that is/are not a key
- **Fixing 3NF violations:**
  - New entity type for offending functional dependency
  - Original entity type:
    - Remove dependent attributes
    - Determinant implements relationship



$A \rightarrow B$ $A \rightarrow C$ $A \rightarrow D$ $A \rightarrow F$

$D \rightarrow F$

$\{B, C\} \rightarrow A$ $\{B, C\} \rightarrow D$ $\{B, C\} \rightarrow F$

Finding functional dependencies, part 4

- **Possible violations of Third Normal Form:**
  - Dependency of a non-key attribute on a non-key attribute
  - Dependency of a non-key attribute on a combination of attributes
    - Not subset of any candidate key
    - Not superset of any candidate key

- **Testing for:**
  - Dependency of ***any*** attribute on a non-key attribute
    - Can be skipped for tables with no non-key attributes
  - Dependency of ***any*** attribute on a combination of attributes
    - Can be skipped for tables with:
      - No composite keys and ≤1 non-key attribute
      - One composite key and no non-key attributes

Finding functional dependencies, part 4

- **Dependencies on a single non-key attribute**
  - Pattern to be populated:
    - One non-key column equal
    - All other columns different
  - **Valid** example with required pattern found or created?
    - No dependencies on non-key column being tested
  - Impossible to create **valid** example with the required pattern?
    - One or more other columns depend on column being tested
      - But which one(s)?

Column A	Column B	Column C	Column D	Column E	Column F	Column G
a1	b1	c1	d1	e1	f1	g1
a2	b2	c2	d1	e2	f2	g2

Finding functional dependencies, part 4

- **Dependencies on a single non-key attribute**
  - Impossible to create **valid** example with the required pattern?
    - One or more other columns depend on column being tested
      - But which one(s)?
    - Check which other columns depend on column being tested
      - Listen to reason for rejecting example
      - Or test, one possibly dependent column at a time

Column A	Column B	Column C	Column D	Column E	Column F	Column G
a1	?	?	d1	?	?	?
a2	?	?	d1	?	?	?

Finding functional dependencies, part 4

- **Dependencies on a single non-key attribute**
  - Impossible to create **valid** example with the required pattern?
    - One or more other columns depend on column being tested
      - But which one(s)?
    - Check which other columns depend on column being tested
      - Listen to reason for rejecting example
      - Or test, one possibly dependent column at a time

Column A	Column B	Column C	Column D	Column E	Column F	Column G
?	b1	?	d1	?	?	?
?	b2	?	d1	?	?	?

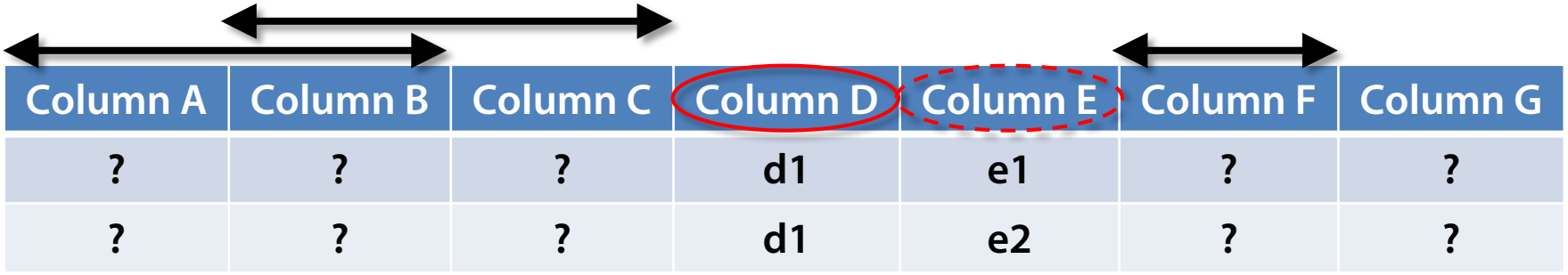
Finding functional dependencies, part 4

- **Dependencies on a single non-key attribute**
  - Impossible to create **valid** example with the required pattern?
    - One or more other columns depend on column being tested
      - But which one(s)?
    - Check which other columns depend on column being tested
      - Listen to reason for rejecting example
      - Or test, one possibly dependent column at a time

Column A	Column B	Column C	Column D	Column E	Column F	Column G
?	?	c1	d1	?	?	?
?	?	c2	d1	?	?	?

Finding functional dependencies, part 4

- **Dependencies on a single non-key attribute**
  - Impossible to create **valid** example with the required pattern?
    - One or more other columns depend on column being tested
      - But which one(s)?
    - Check which other columns depend on column being tested
      - Listen to reason for rejecting example
      - Or test, one possibly dependent column at a time



Column A	Column B	Column C	Column D	Column E	Column F	Column G
?	?	?	d1	e1	?	?
?	?	?	d1	e2	?	?

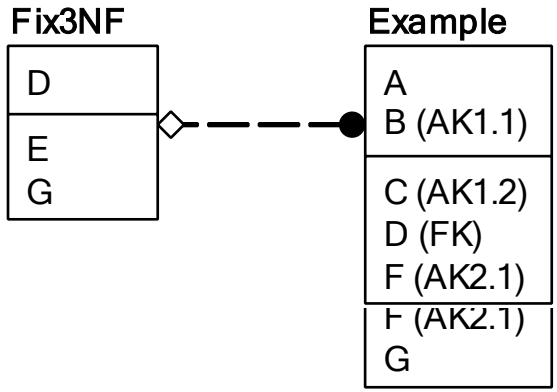
Finding functional dependencies, part 4

- **Dependencies on a single non-key attribute**
  - Impossible to create **valid** example with the required pattern?
    - One or more other columns depend on column being tested
      - But which one(s)?
    - Check which other columns depend on column being tested
      - Listen to reason for rejecting example
      - Or test, one possibly dependent column at a time
      - When rejected, that column depends on column being tested
        - (Usually)
      - Make sure to find *all* dependent columns!
    - For each dependency: "normal" functional dependency, or derivation rule?

Column A	Column B	Column C	Column D	Column E	Column F	Column G
?	?	?	d1	?	?	g1
?	?	?	d1	?	?	g2

Finding functional dependencies, part 4

- **Dependencies on a single non-key attribute**
  - Impossible to create **valid** example with the required pattern?
    - "Normal" (not derived) dependencies of non-key column → violates 3NF
      - Create new entity type
      - Remove dependent attribute(s)



Functional dependencies found: $D \rightarrow E$

$D \rightarrow G$

Column A	Column B	Column C	Column D	Column E	Column F	Column G

Finding functional dependencies, part 4

- **Dependencies on a single non-key attribute**
  - Impossible to create **valid** example with the required pattern?
    - Derived dependencies / dependent key columns
      - No schema change needed
      - Make note for later use
      - Exclude dependent column from rest of test
        - But do check for transitive dependencies if new dependency on determinant is found!

Example

A B (AK1.1)
C (AK1.2) D E F (AK2.1) G

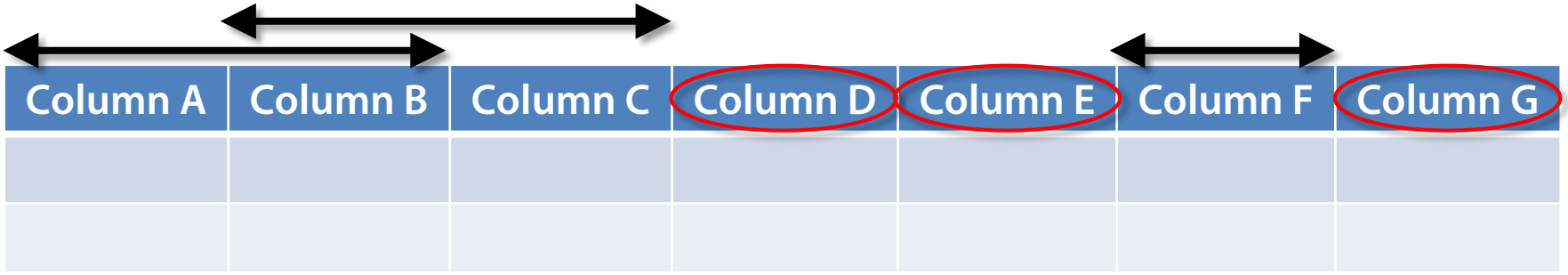
Functional dependencies found: D → E

D → G

Column A	Column B	Column C	Column D	Column E	Column F	Column G

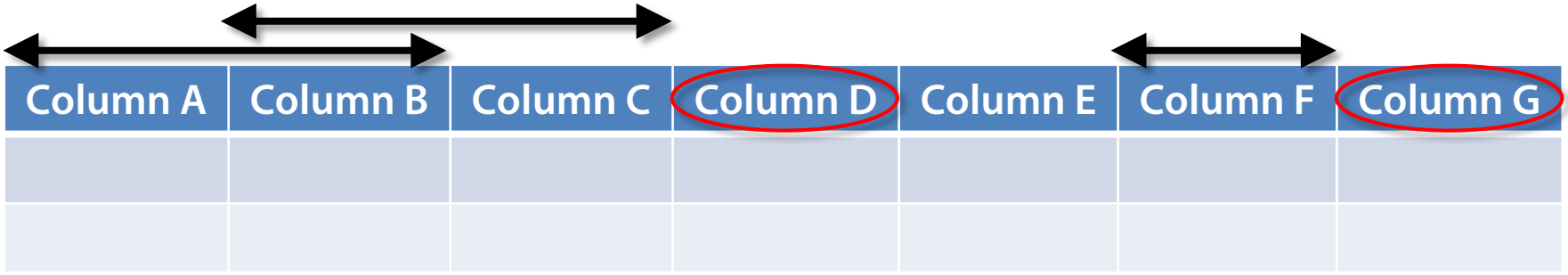
Finding functional dependencies, part 4

- **Dependencies on a single non-key attribute**
  - When tests for a column are completed:
    - **First** fix model (if needed)
    - Then continue with next column
      - Repeat until all columns verified



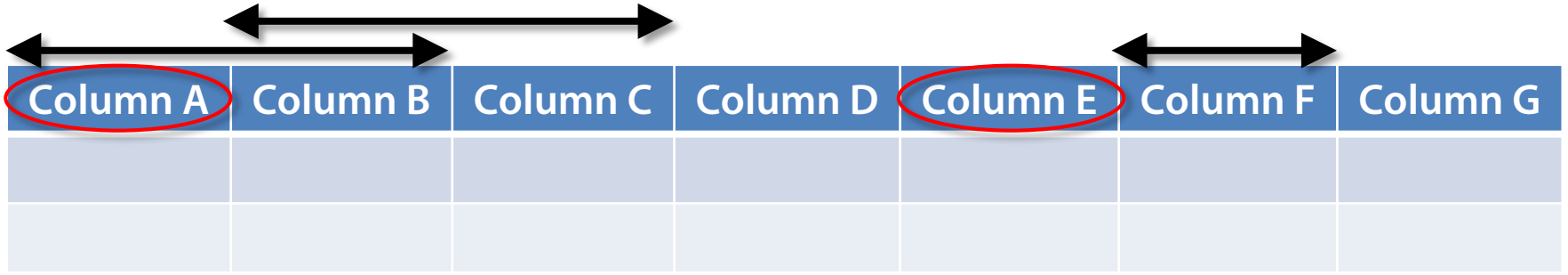
Finding functional dependencies, part 4

- **Dependencies on combination of attributes that are not a key**
  - Test possible combinations of two columns, then three, …
    - Combination of several non-key columns



Finding functional dependencies, part 4

- **Dependencies on combination of attributes that are not a key**
  - Test possible combinations of two columns, then three, …
    - Combination of several non-key columns
    - Combination of non-key column(s) with subset of a candidate key

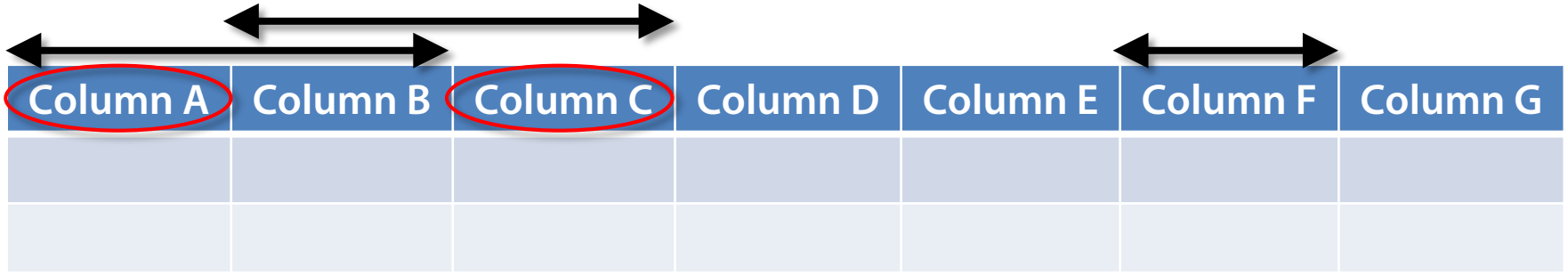


Finding functional dependencies, part 4

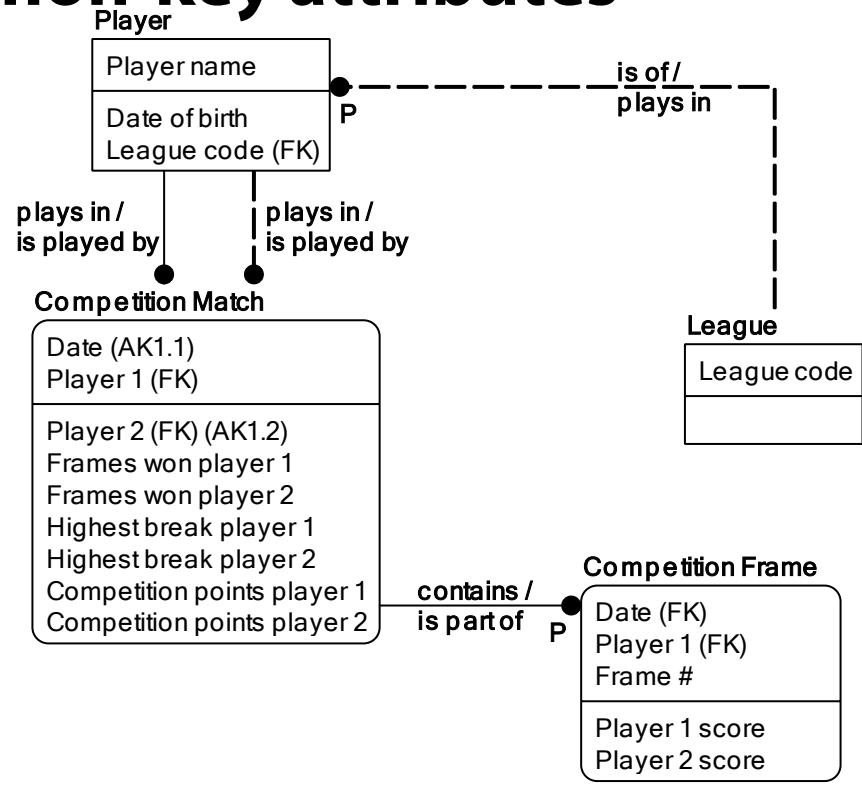
- **Dependencies on combination of attributes that are not a key**
  - Test possible combinations of two columns, then three, …
    - Combination of several non-key columns
    - Combination of non-key column(s) with subset of a candidate key
    - Combination of several subsets of different candidate keys
  - Problem: This results in a LOT of test cases!
    - Solution: use your experience to avoid "silly" tests
    - But remember that this involves assumptions!

Not a subset of a candidate key

Not a superset of a candidate key



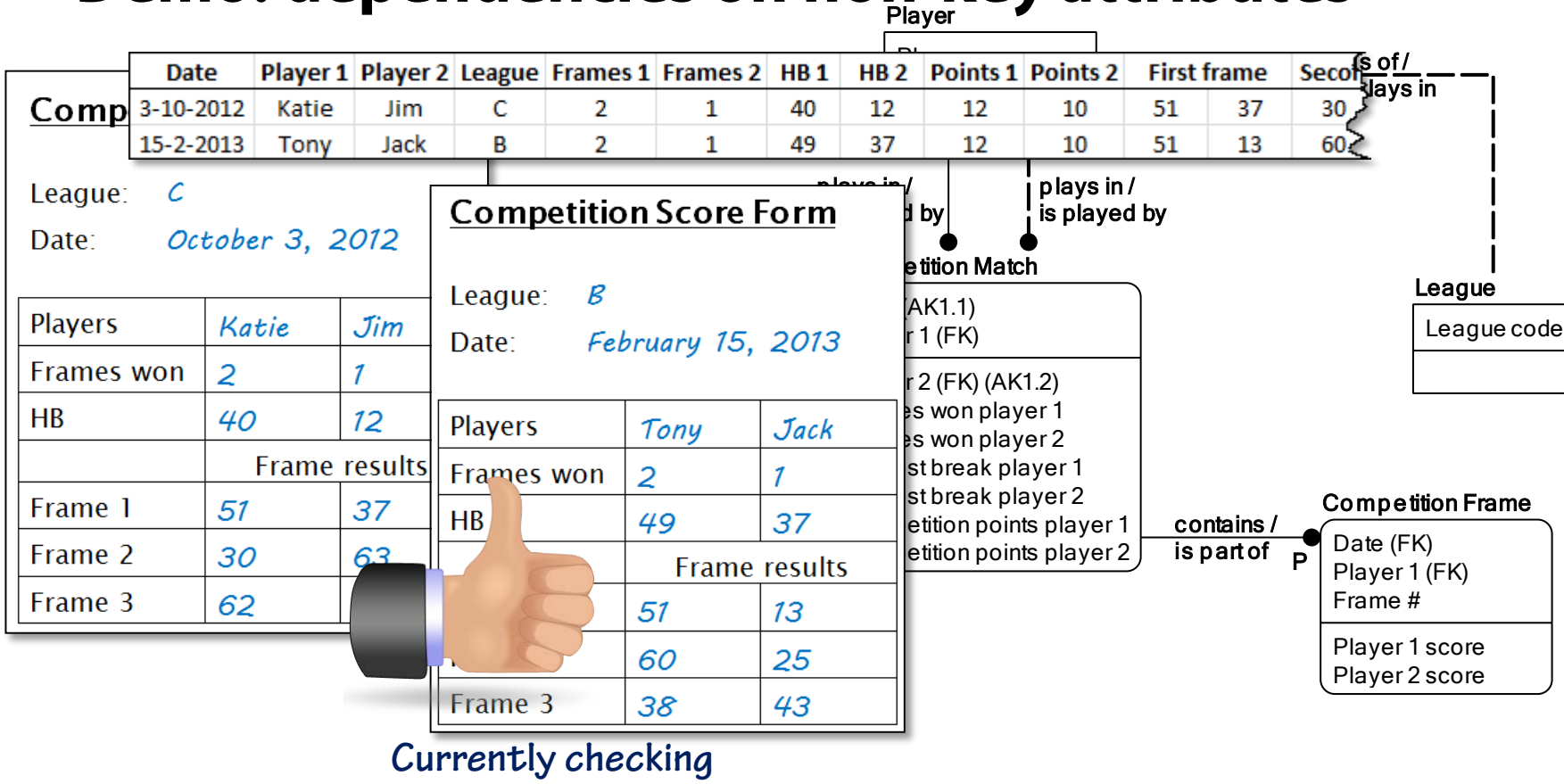
Demo: dependencies on non-key attributes



Currently checking

Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2
a1	b1	c1	d1	e1	f1	g1	h1	j1
a2	b2	c2	d1	e2	f2	g2	h2	j2

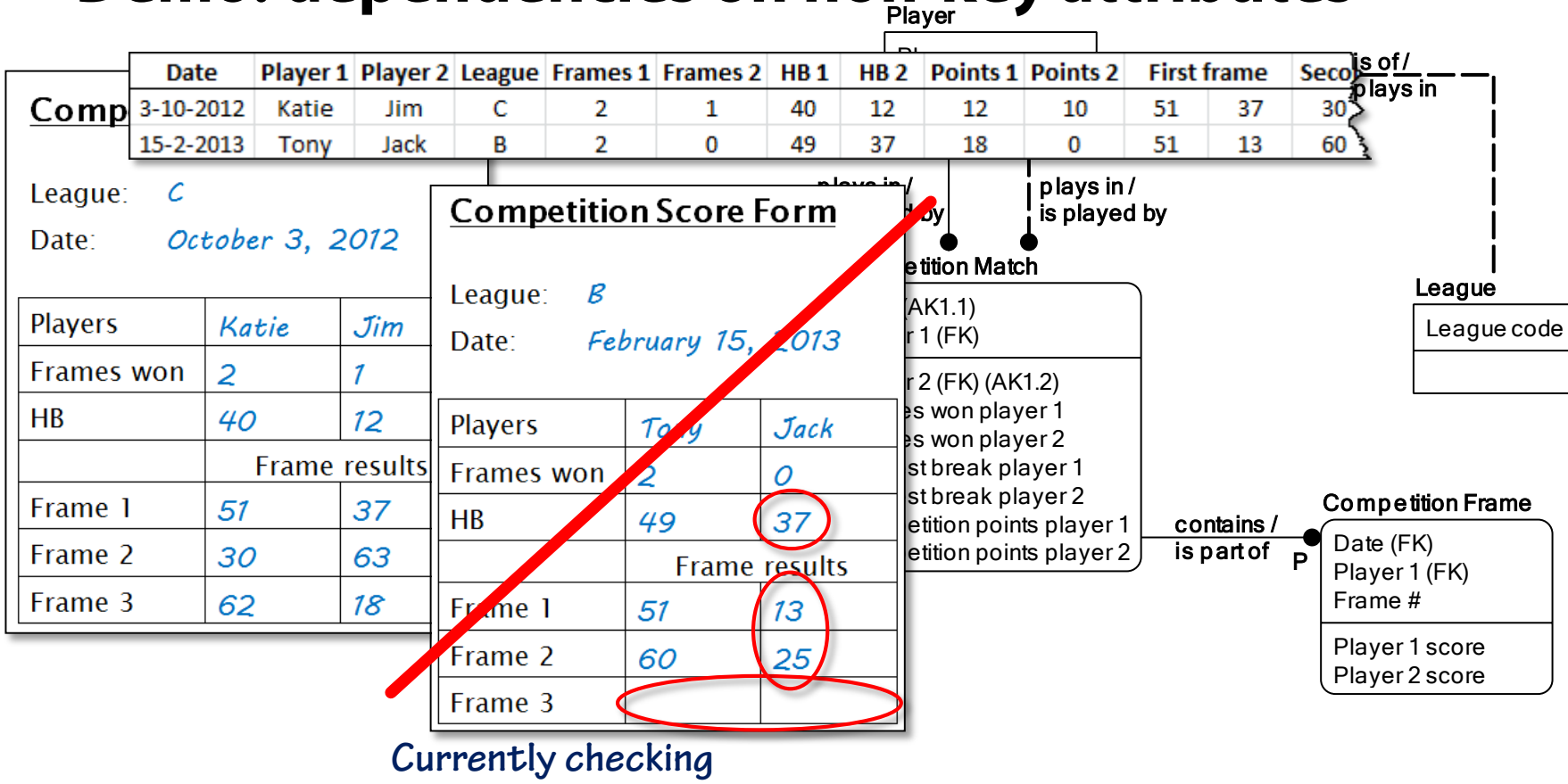
Demo: dependencies on non-key attributes



Currently checking

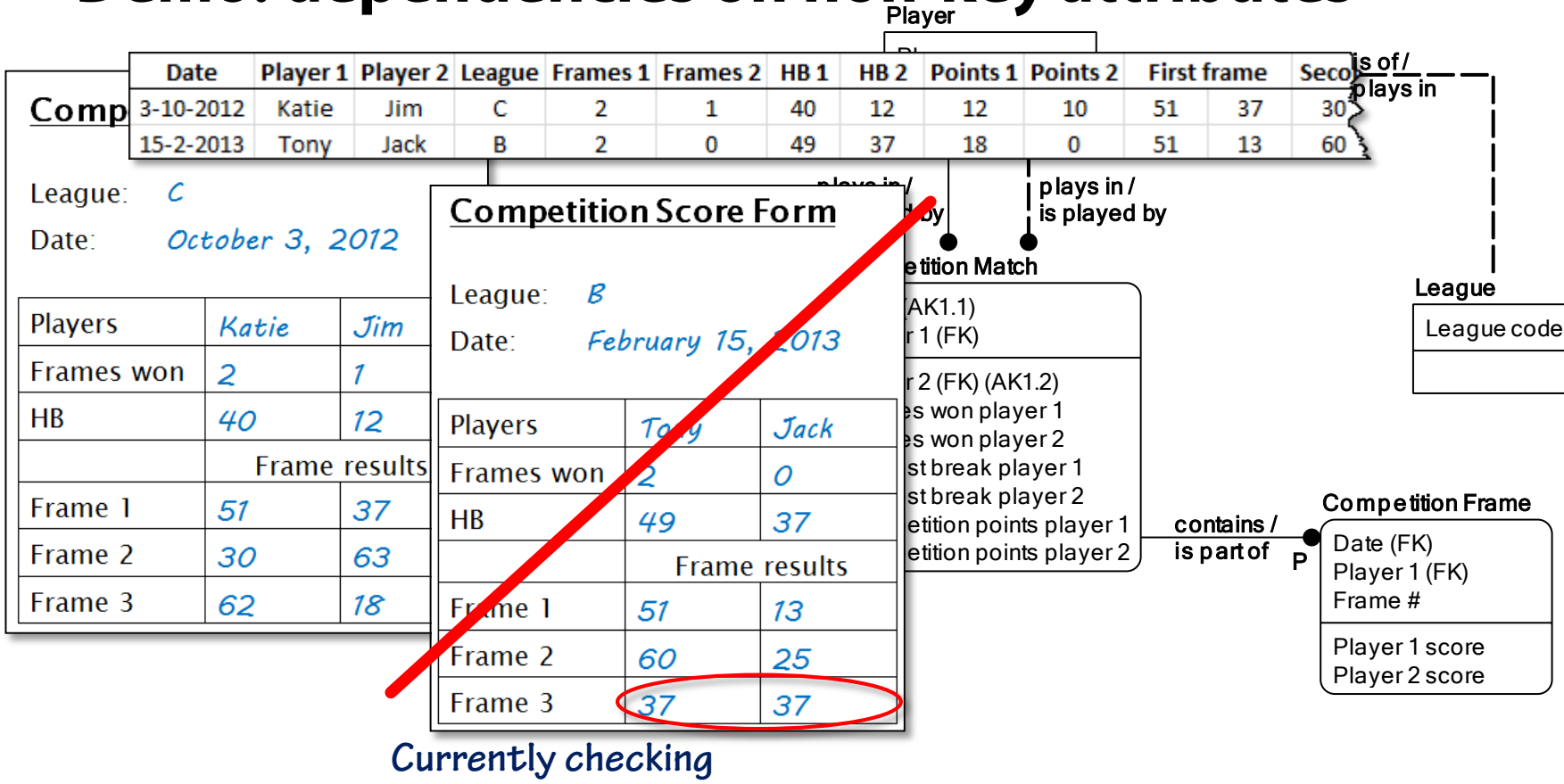
Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2
2012-10-03	Katie	Jim	2	1	40	12	12	10
2013-02-15	Tony	Jack	2	1	49	37	12	10

Demo: dependencies on non-key attributes



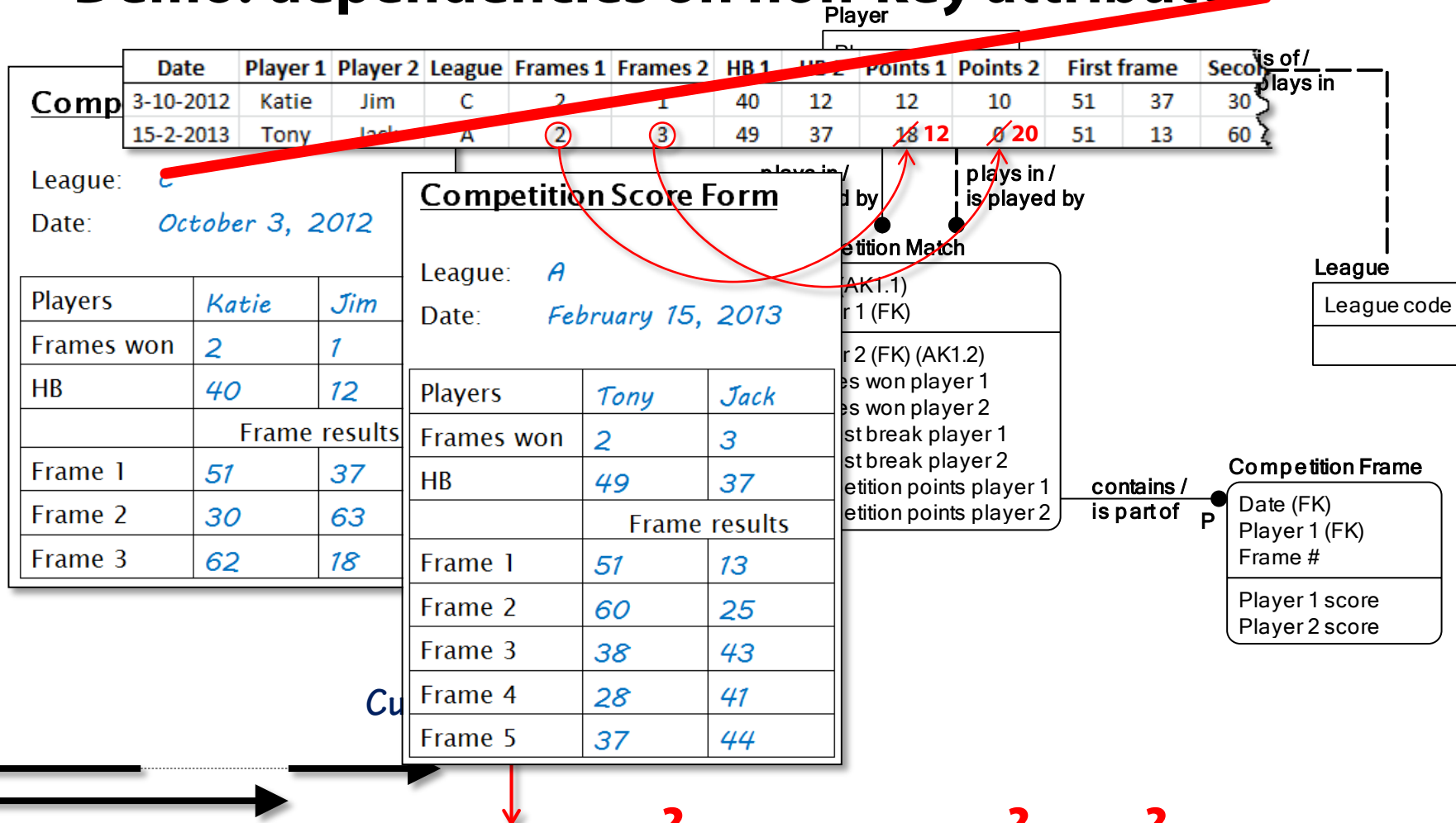
Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2
2012-10-03	Katie	Jim	2	1	40	12	12	10
2013-02-15	Tony	Jack	2	0	49	37	18	0

Demo: dependencies on non-key attributes



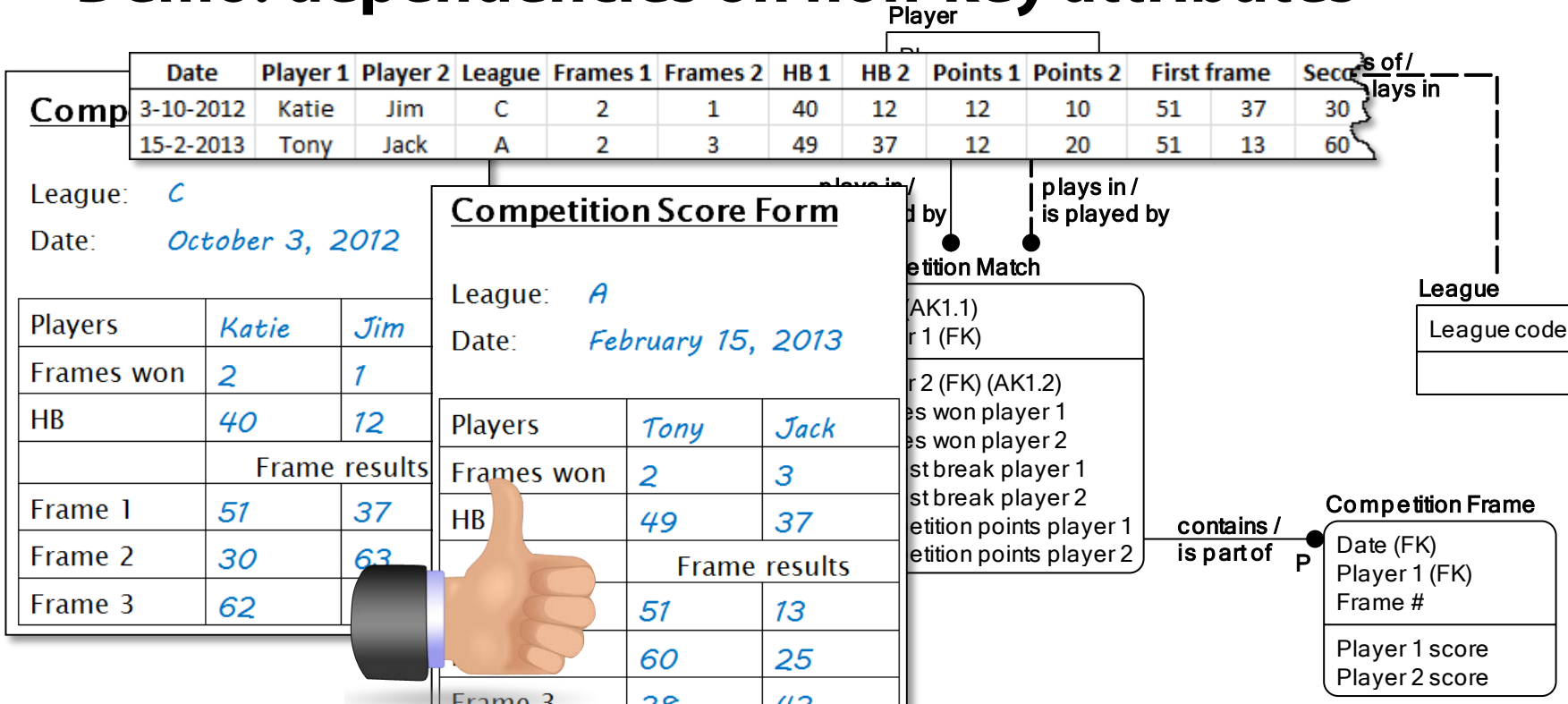
Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2
2012-10-03	Katie	Jim	2	1	40	12	12	10
2013-02-15	Tony	Jack	2	0	49	37	18	0

Demo: dependencies on non-key attributes



Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2
2012-10-03	Katie	Jim	2	1	40	12	12	10
2013-02-15	Tony	Jack	2	3	49	37	18	0

Demo: dependencies on non-key attributes

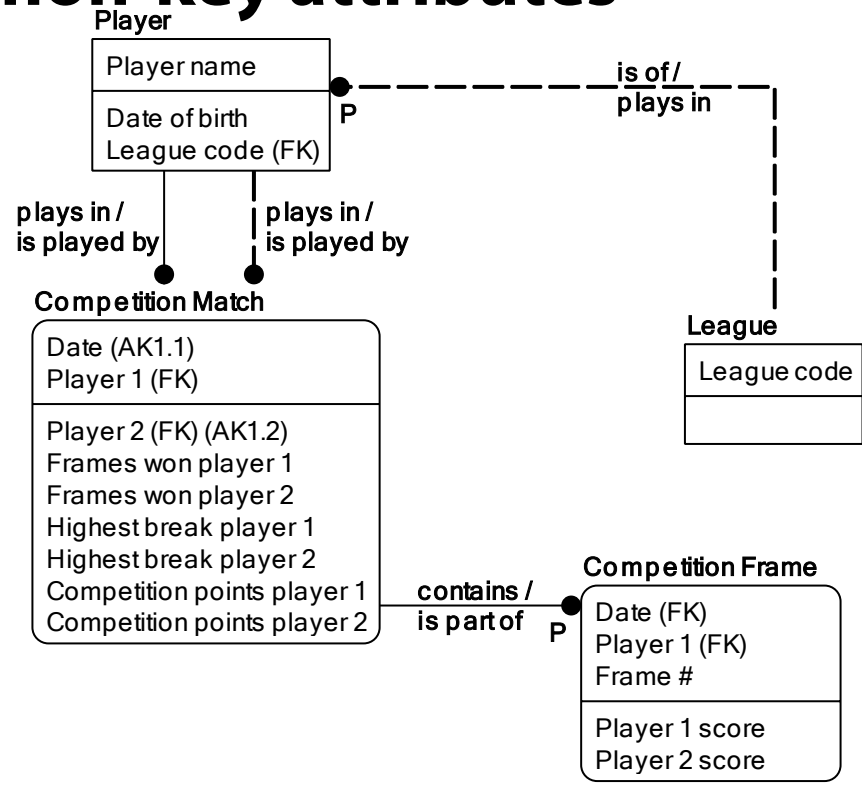


Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2
2012-10-03	Katie	Jim	2	1	40	12	12	10
2013-02-15	Tony	Jack	2	3	49	37	12	20

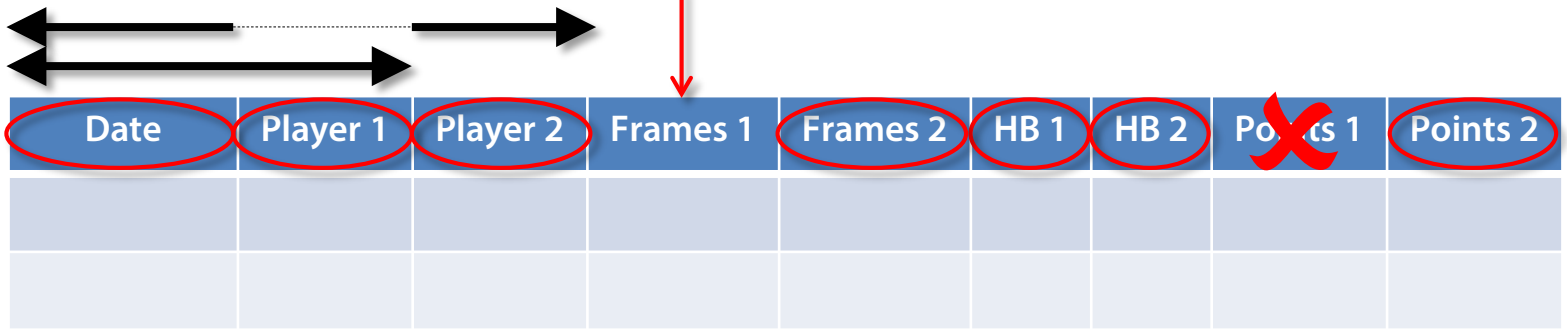
Demo: dependencies on non-key attributes

Dependency found:
- Frames 1 → Points 1

Functional dependency or derivation rule?



Currently checking

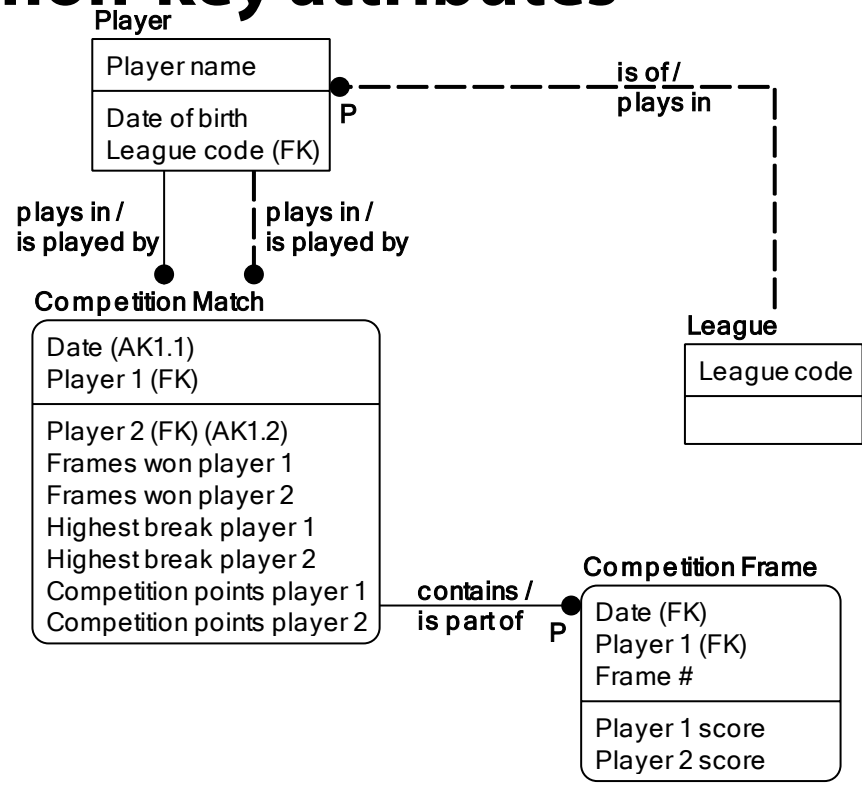


Demo: dependencies on non-key attributes

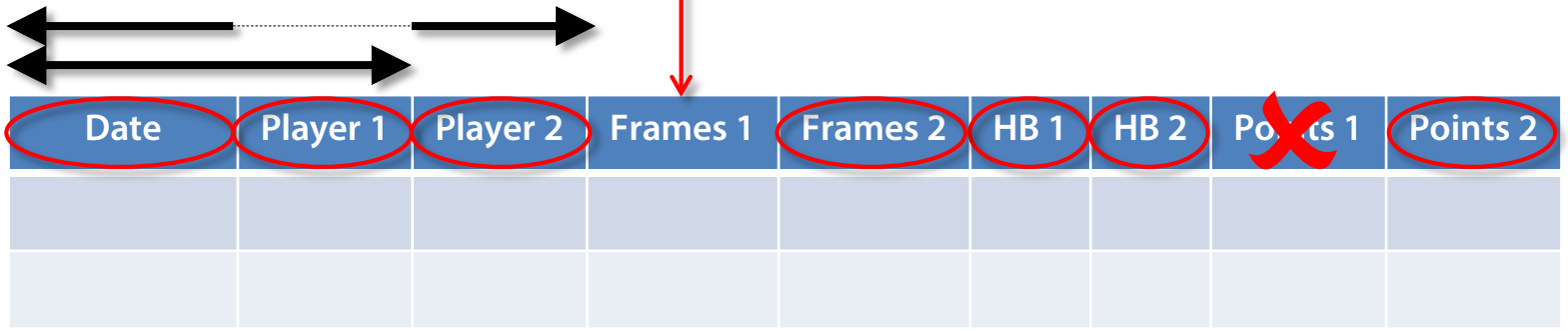
Dependency found:
- Frames 1 → Points 1

Functional dependency!

Violates 3NF!



Currently checking

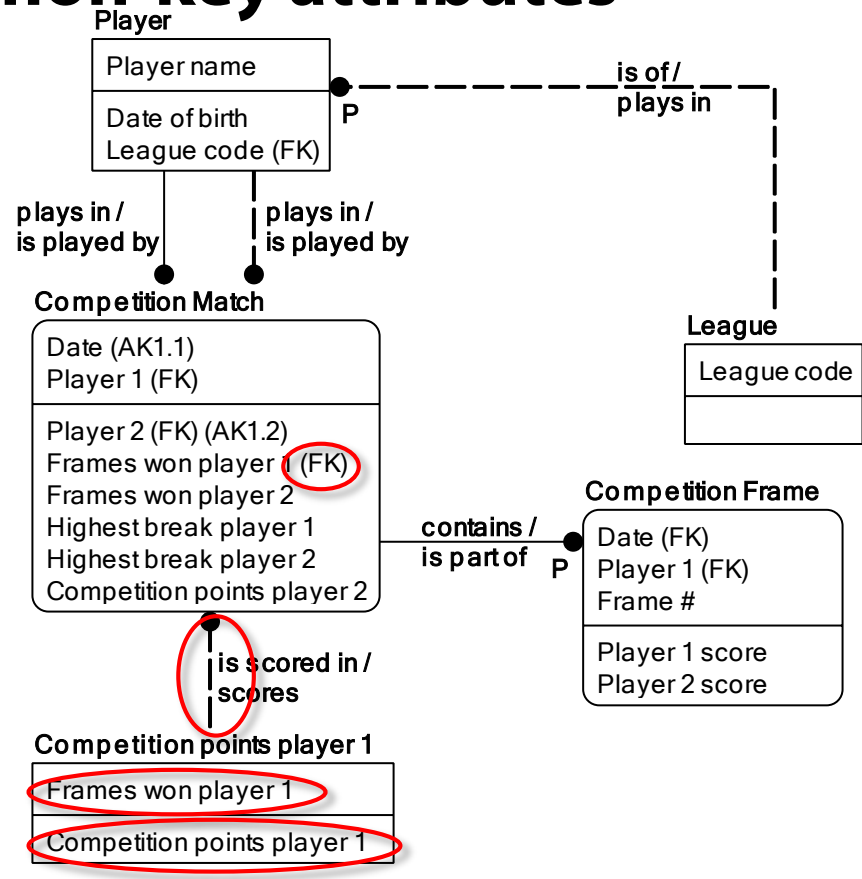


Demo: dependencies on non-key attributes

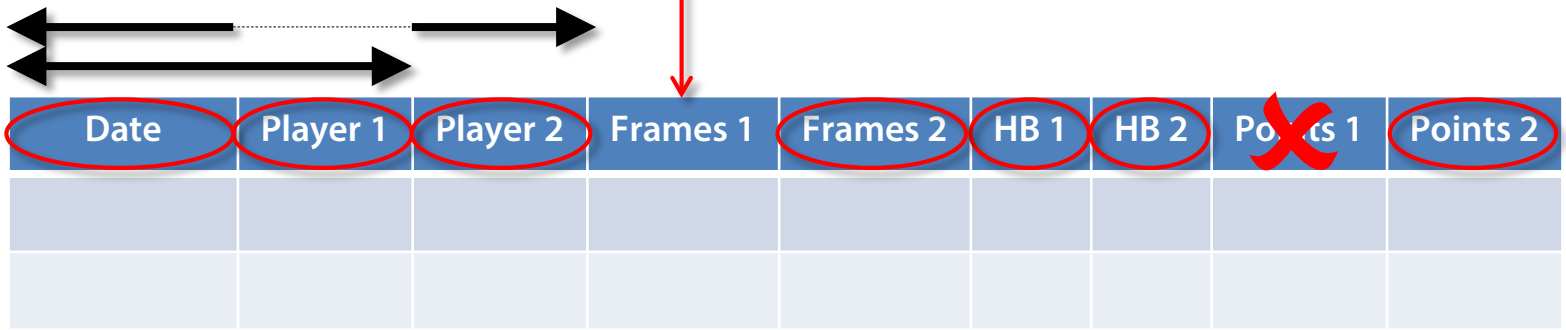
Dependency found:
- Frames 1 → Points 1

Functional dependency!

Violates 3NF!



Currently checking



Demo: dependencies on non-key attributes

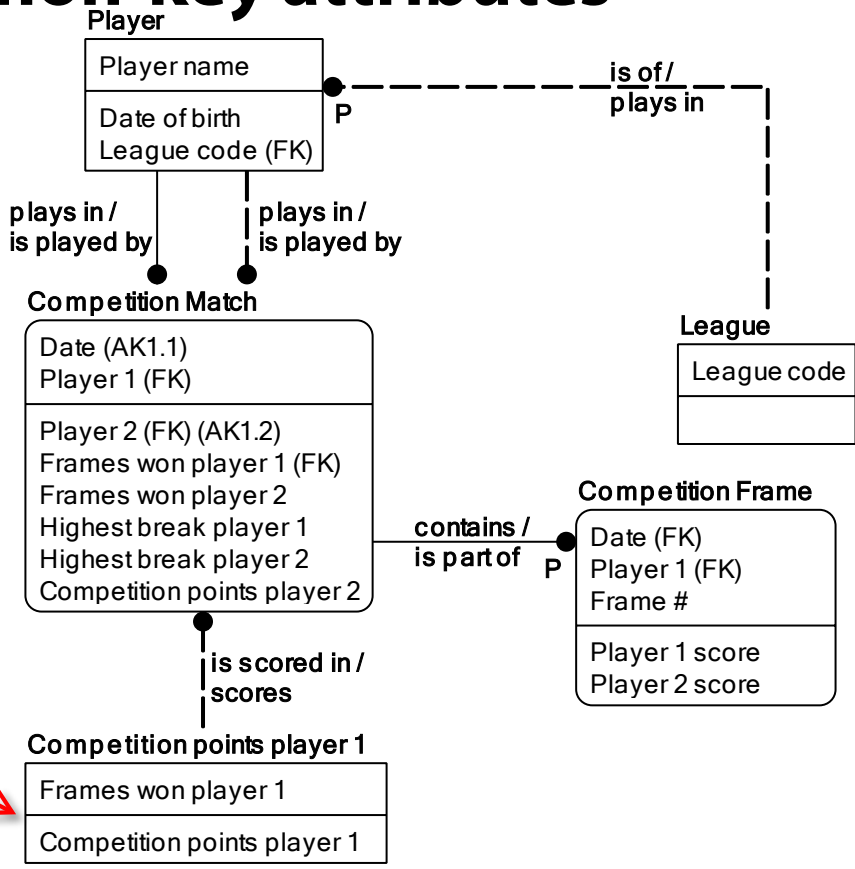
Dependency found:
- Frames 1 → Points 1

Functional dependency!

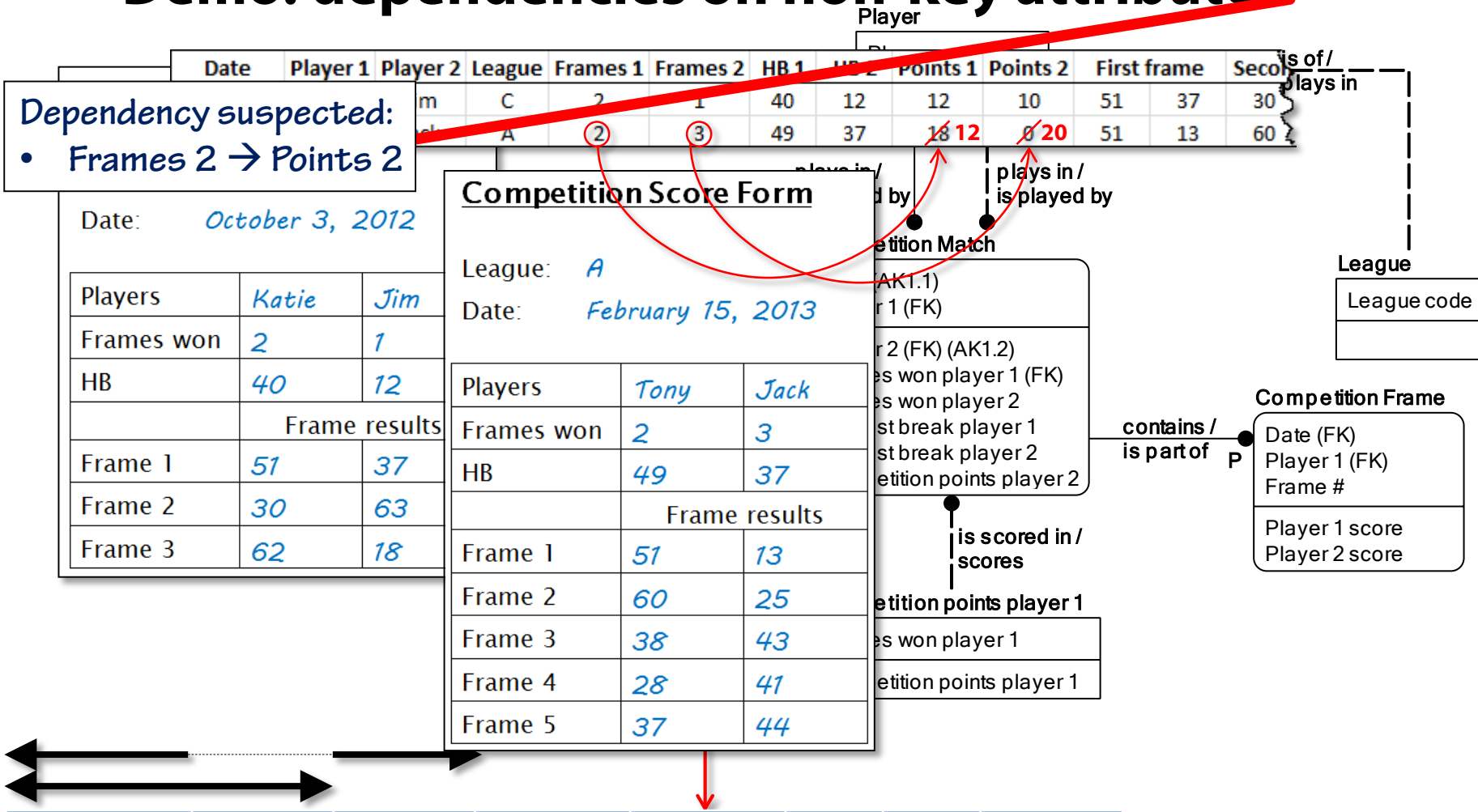
Frames won	Competition points
0	0
1	10
2	12
3	20
4	30
5	45

Currently checking

Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2	Points 1	Points 2



Demo: dependencies on non-key attributes



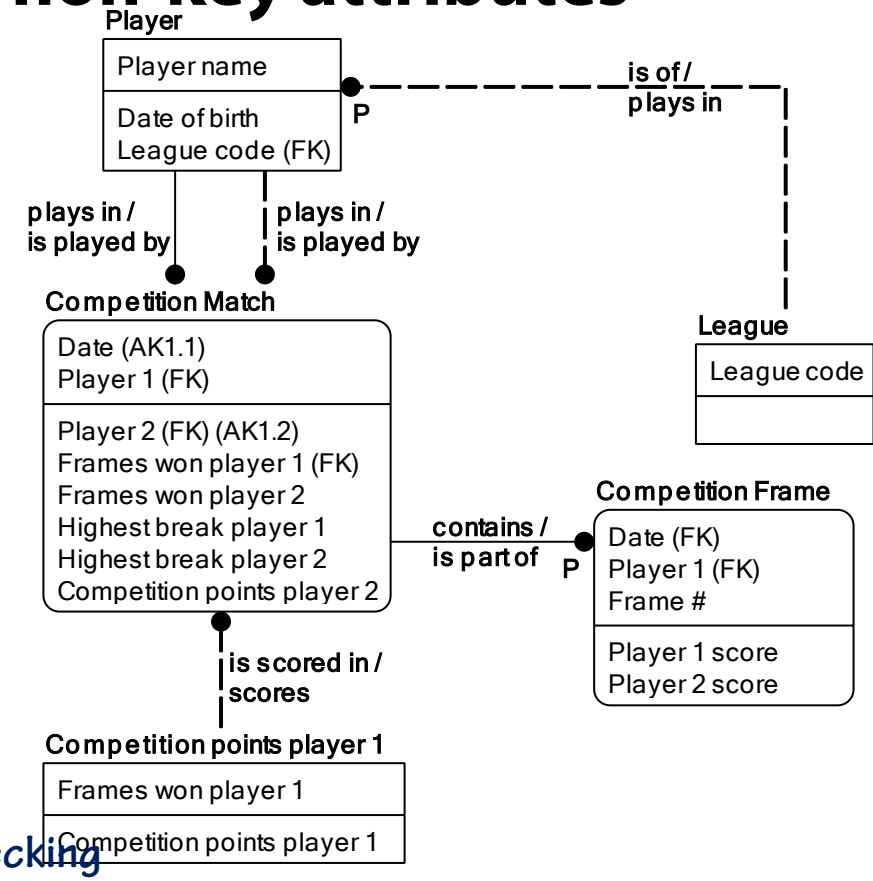
Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2	Points 2
2012-10-03	Katie	Jim	2	1	40	12	10
2013-02-15	Tony	Jack	2	3	49	37	0

Demo: dependencies on non-key attributes

Dependency found:
- Frames 2 → Points 2

Functional dependency!

Violates 3NF!



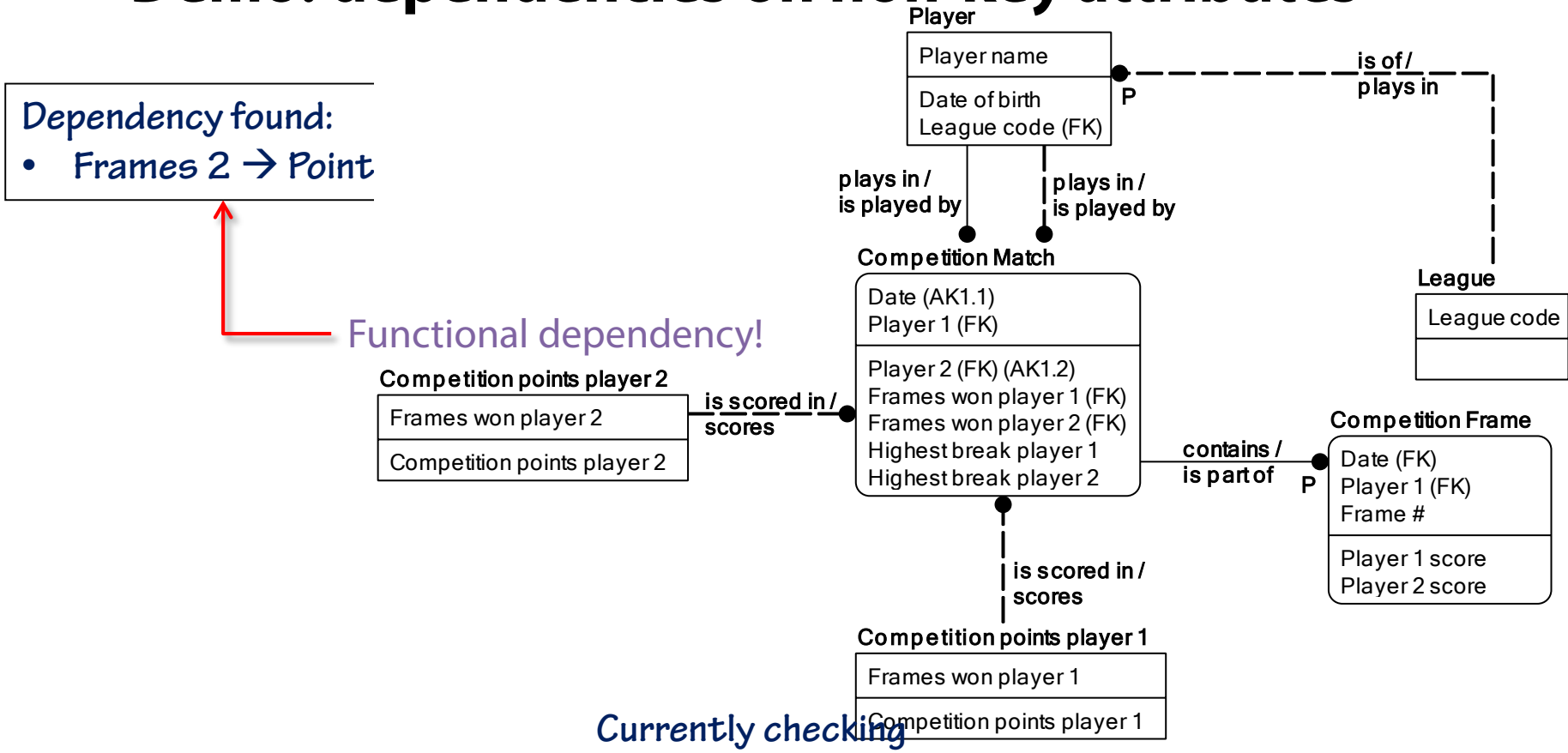
Currently checking

Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2	Points 2

Demo: dependencies on non-key attributes

Dependency found:
- Frames 2 → Point

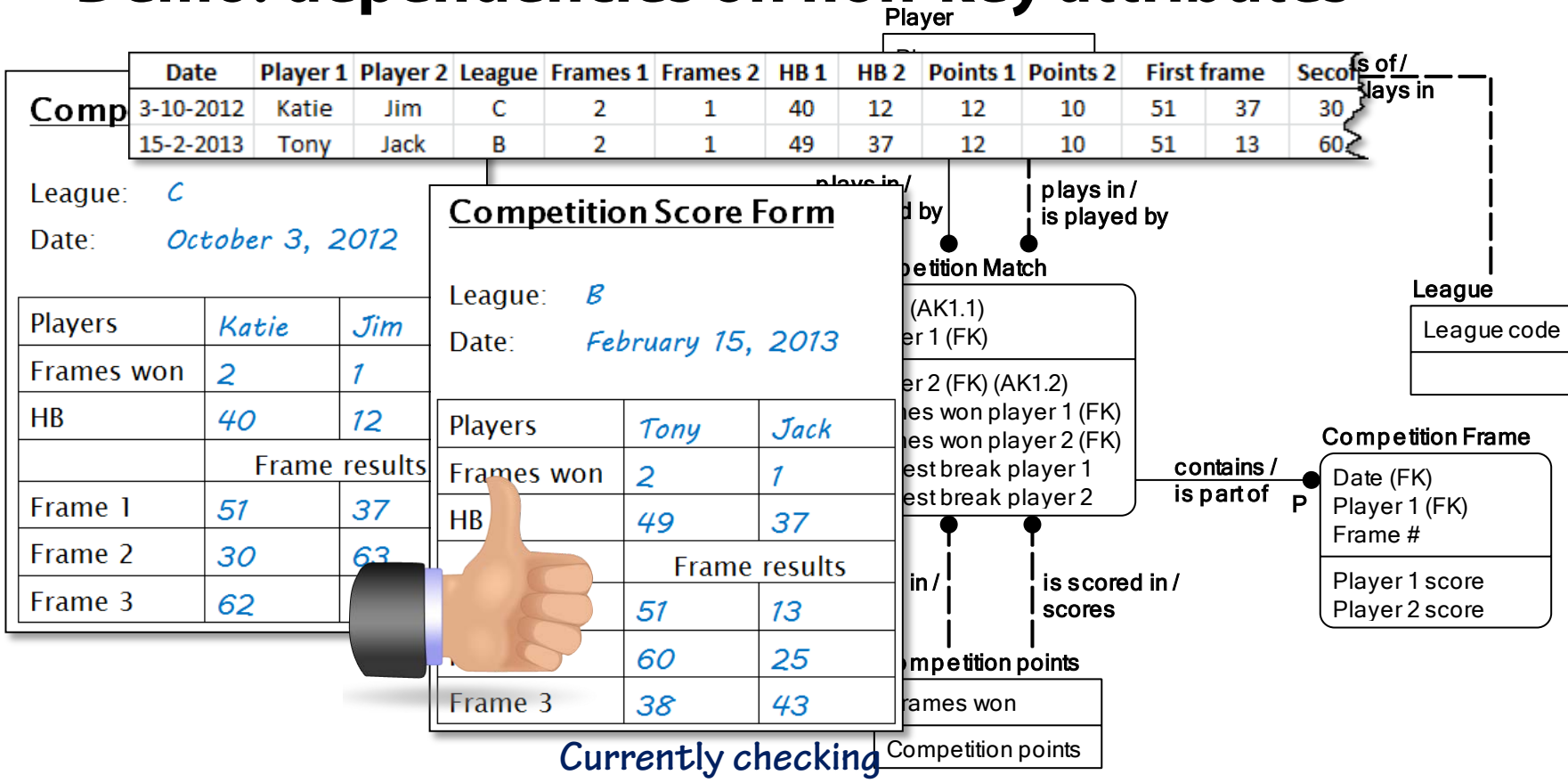
Functional dependency!



Currently checking

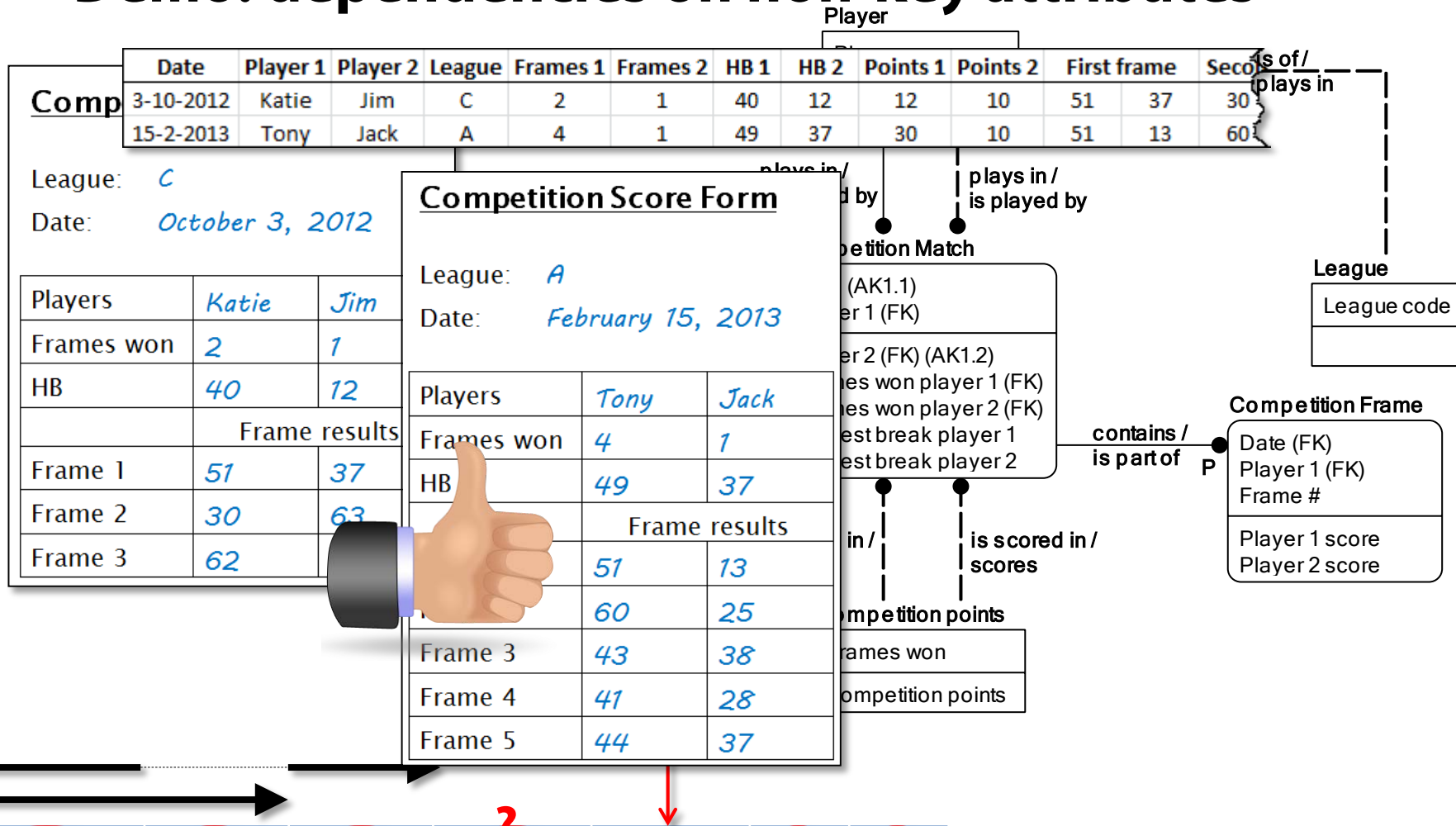
Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2	Points 2

Demo: dependencies on non-key attributes

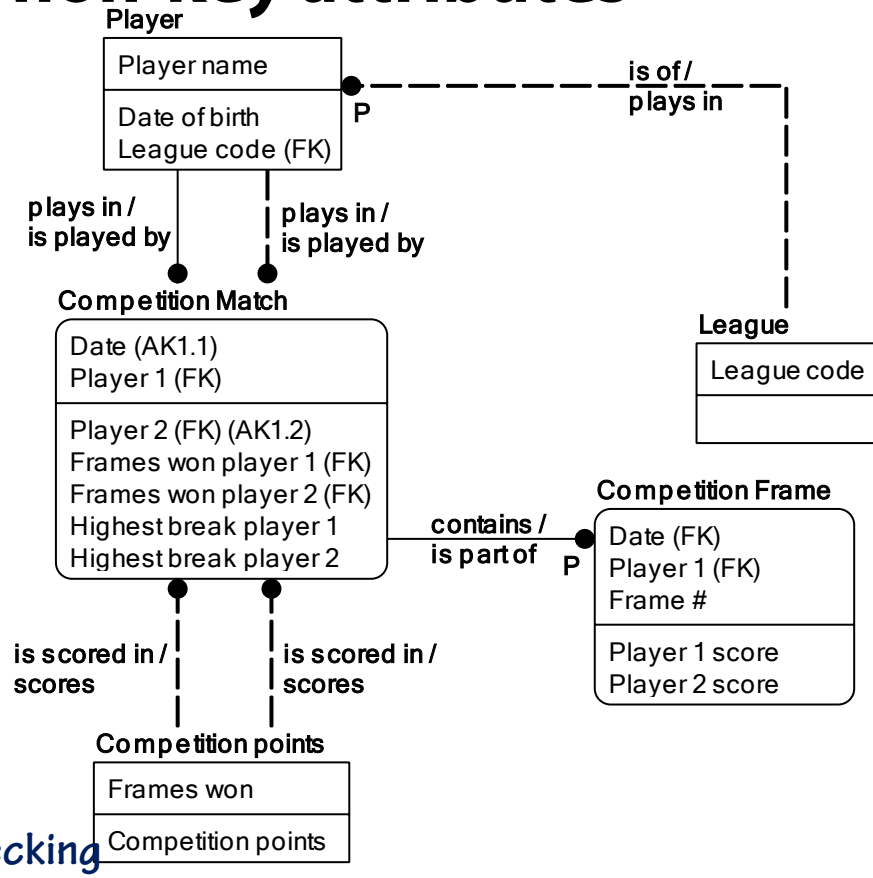


Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2
2012-10-03	Katie	Jim	2	1	40	12
2013-02-15	Tony	Jack	2	1	49	37

Demo: dependencies on non-key attributes



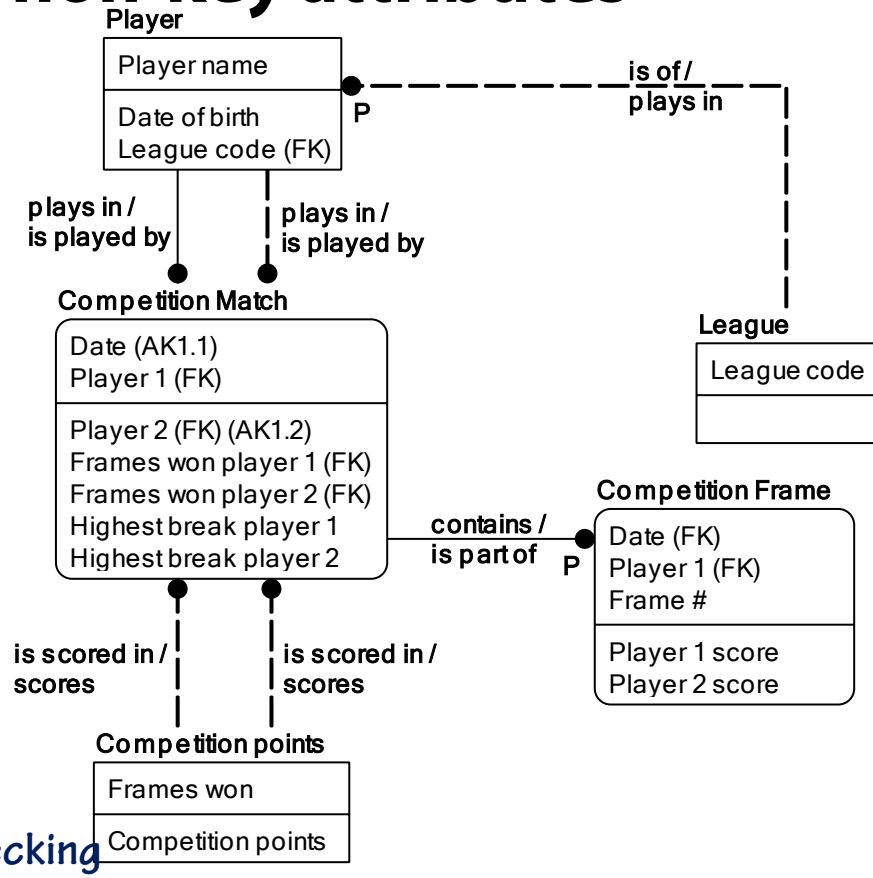
Demo: dependencies on non-key attributes



Currently checking

Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2

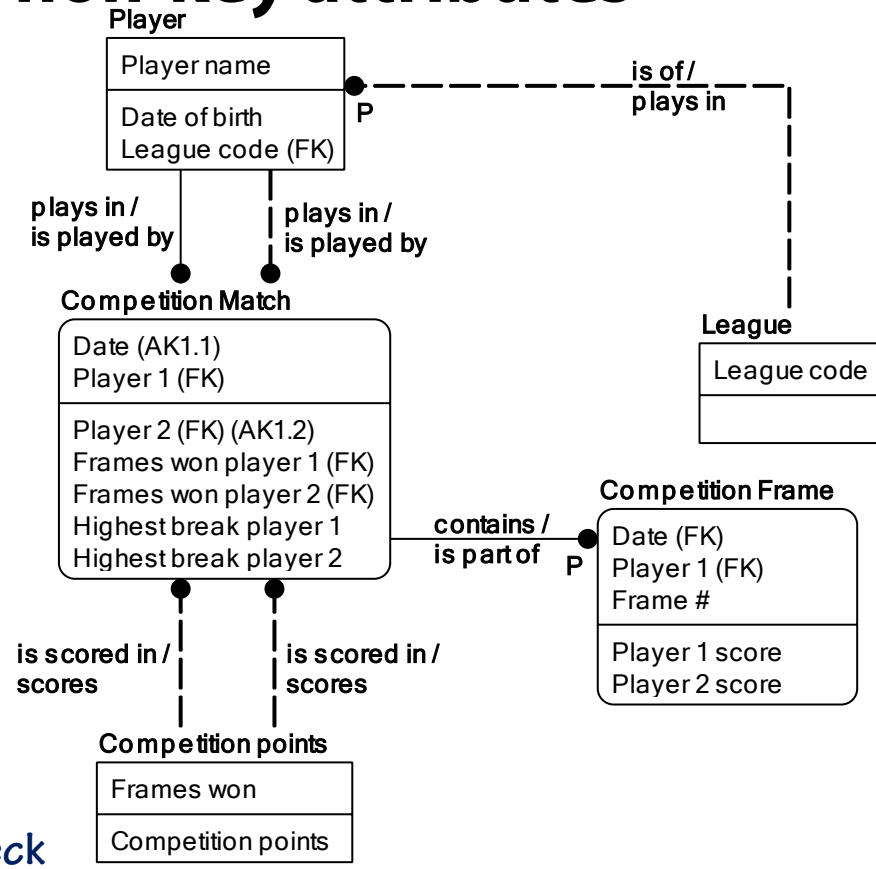
Demo: dependencies on non-key attributes



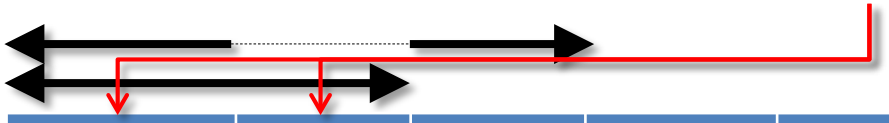
Currently checking

Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2

Demo: dependencies on non-key attributes

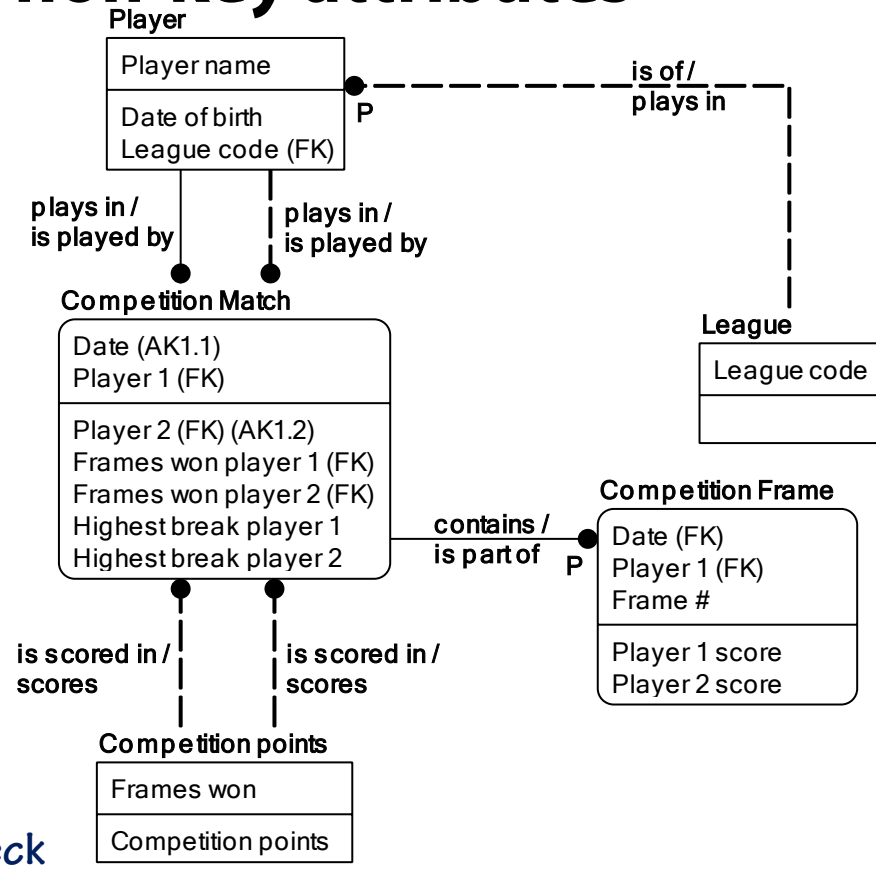


Do not check

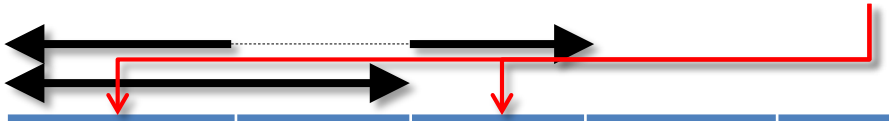


Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2

Demo: dependencies on non-key attributes

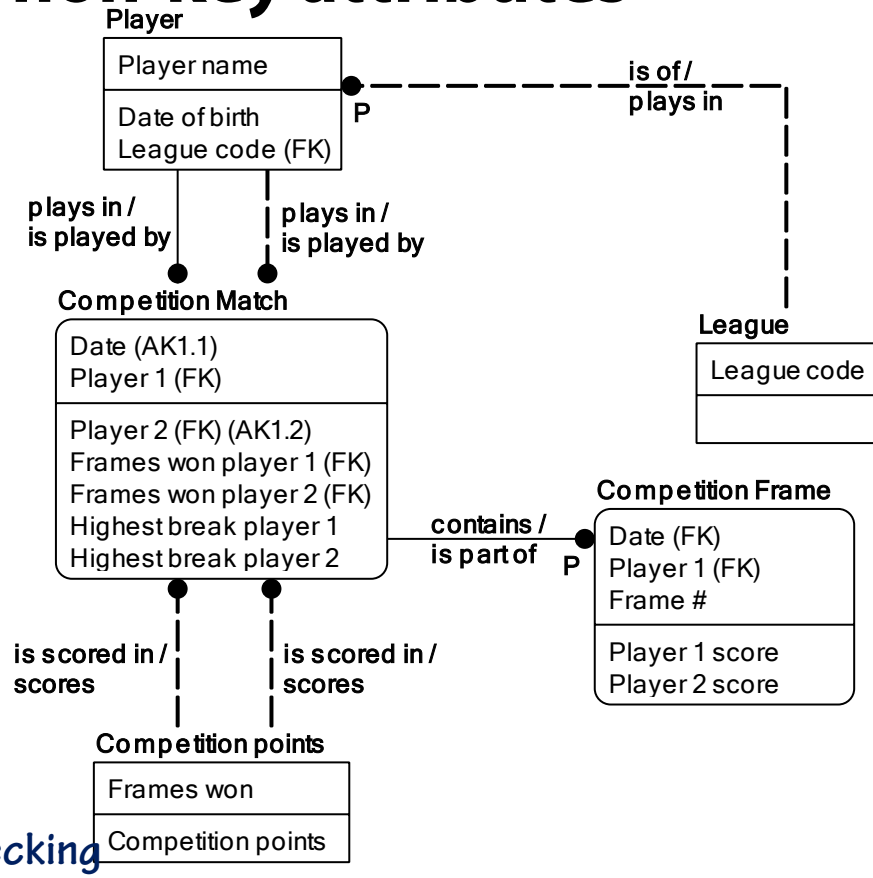


Do not check



Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2

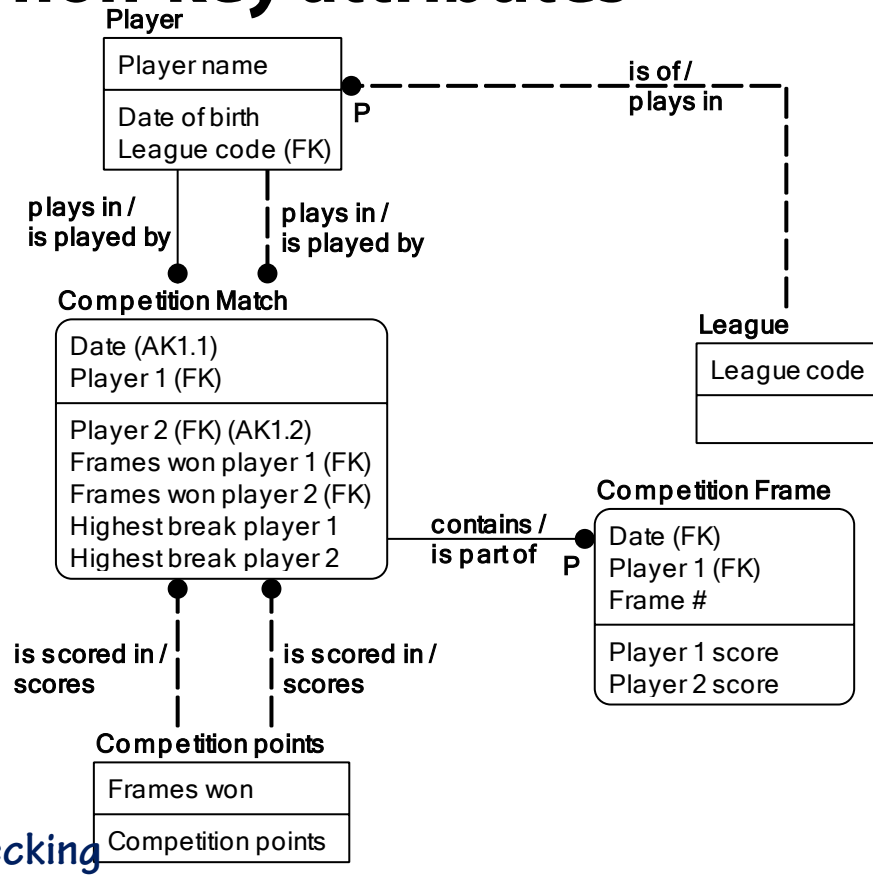
Demo: dependencies on non-key attributes



Currently checking

Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2

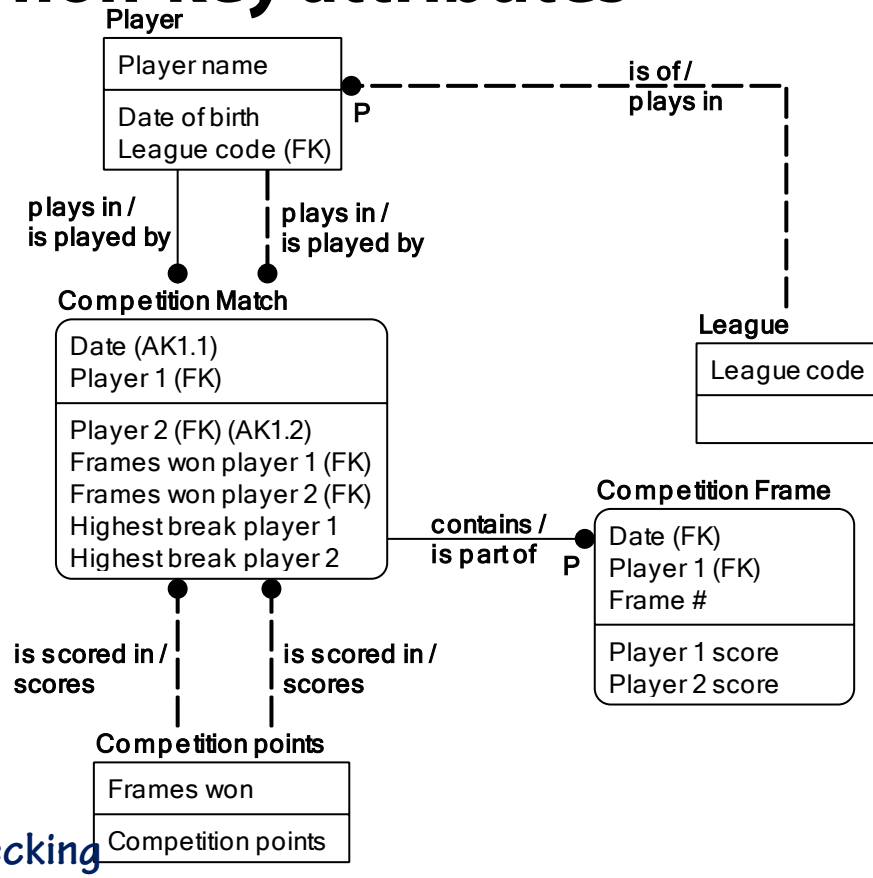
Demo: dependencies on non-key attributes



Currently checking

Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2

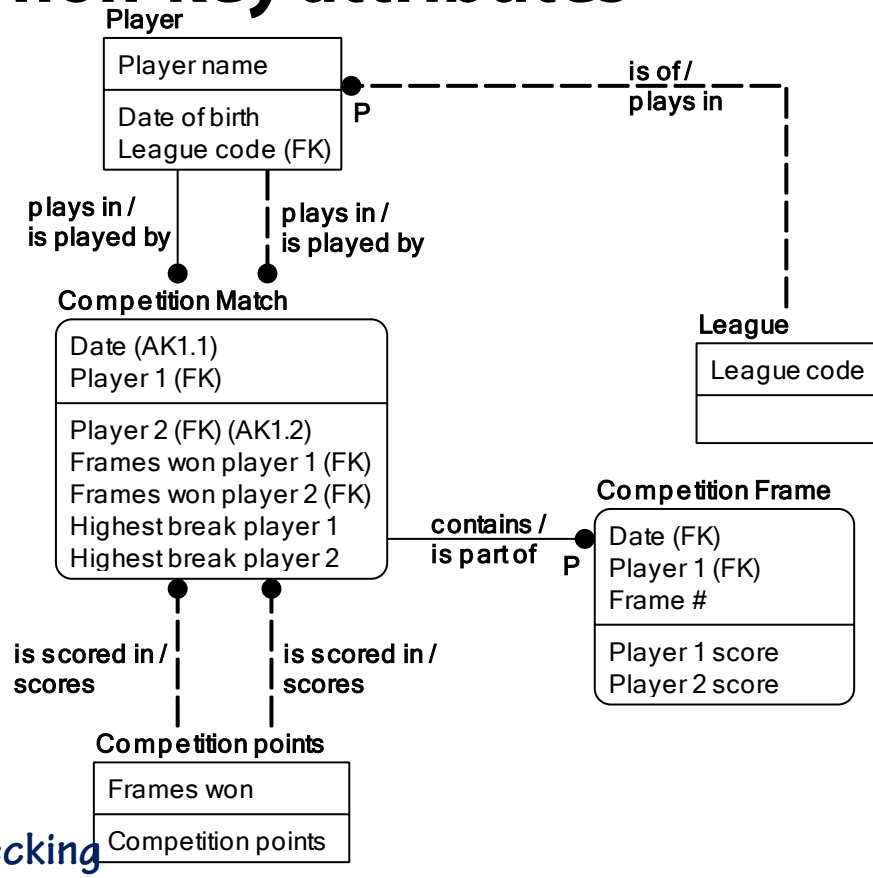
Demo: dependencies on non-key attributes



Currently checking

Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2

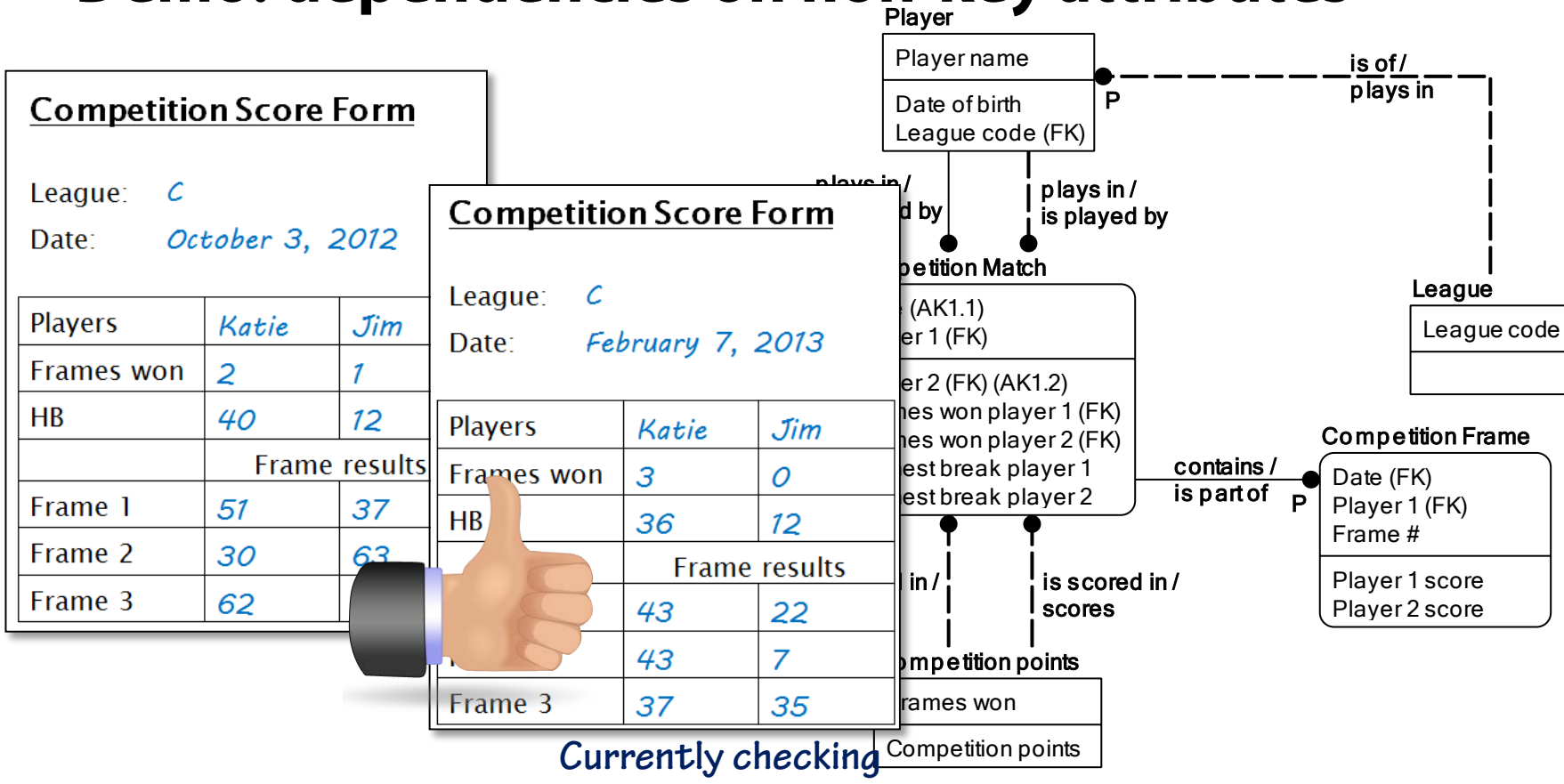
Demo: dependencies on non-key attributes



Currently checking

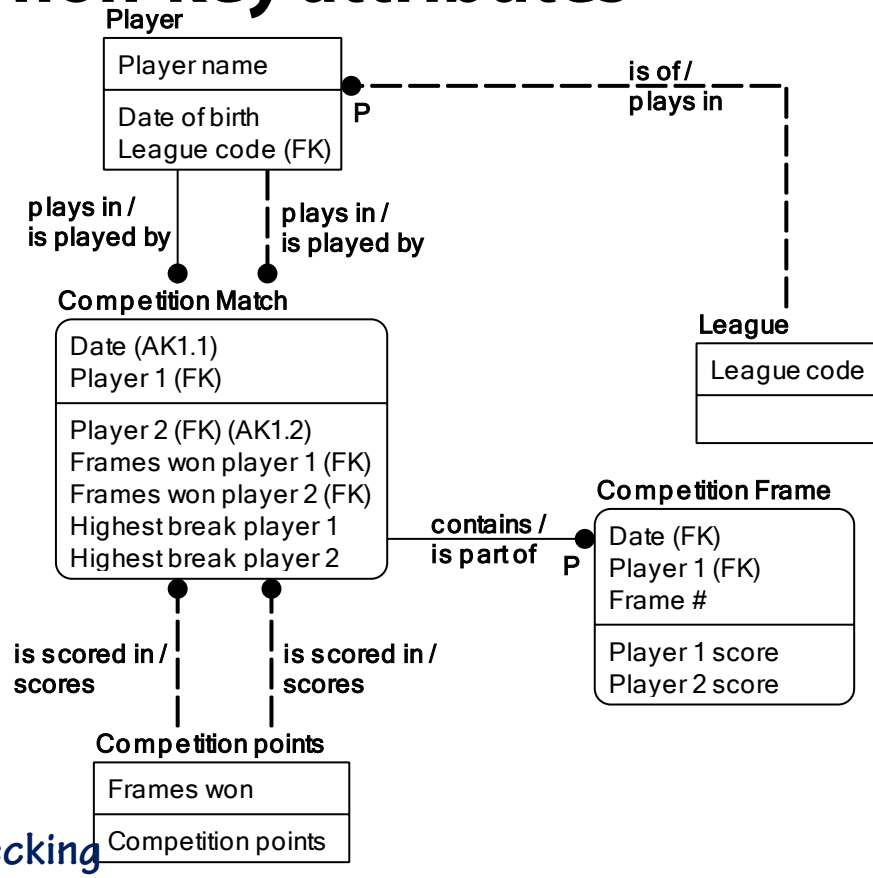
Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2

Demo: dependencies on non-key attributes



Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2
2012-10-03	Katie	Jim	2	1	40	12
2013-02-07	Katie	Jim	3	0	36	12

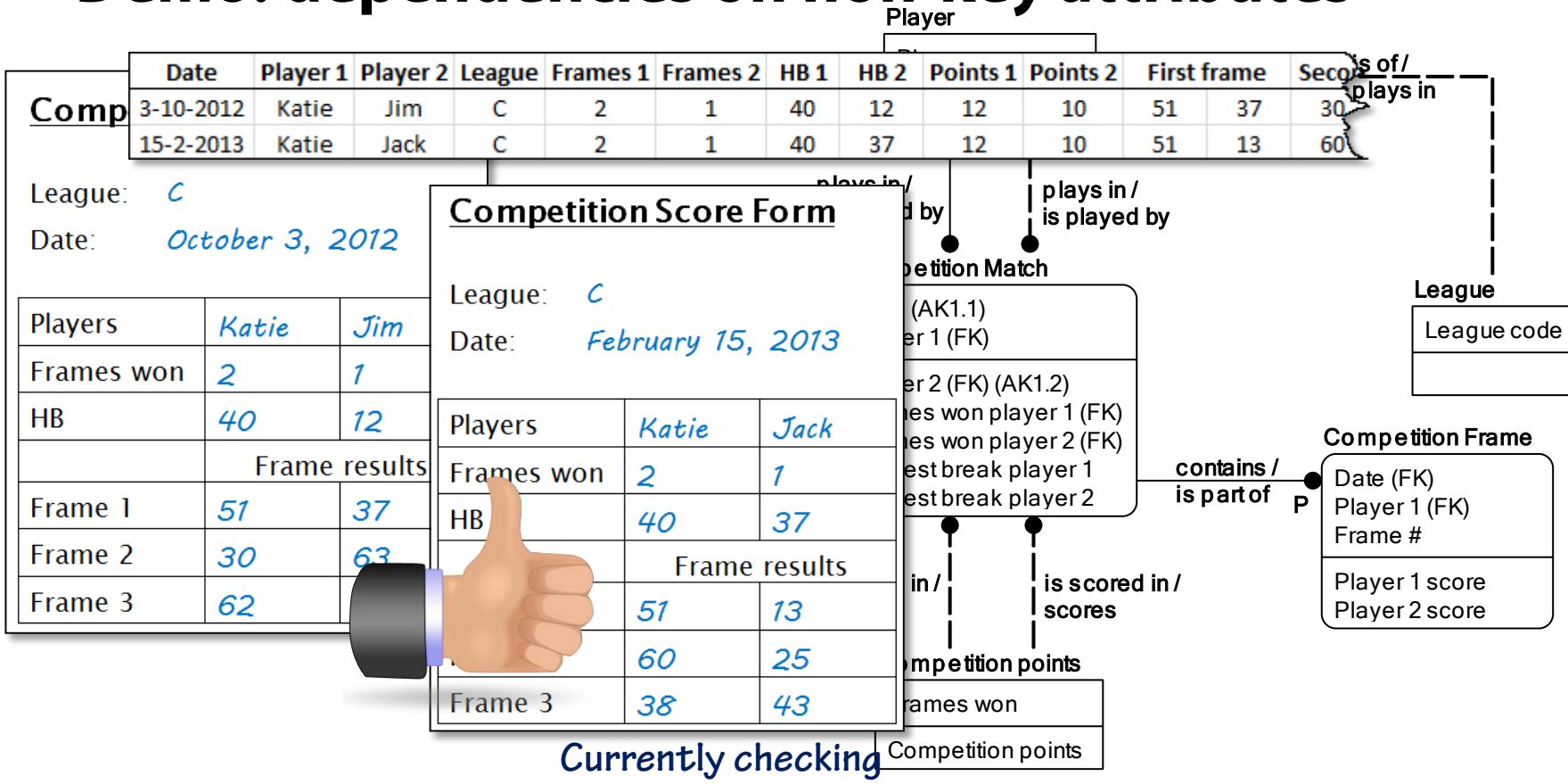
Demo: dependencies on non-key attributes



Currently checking

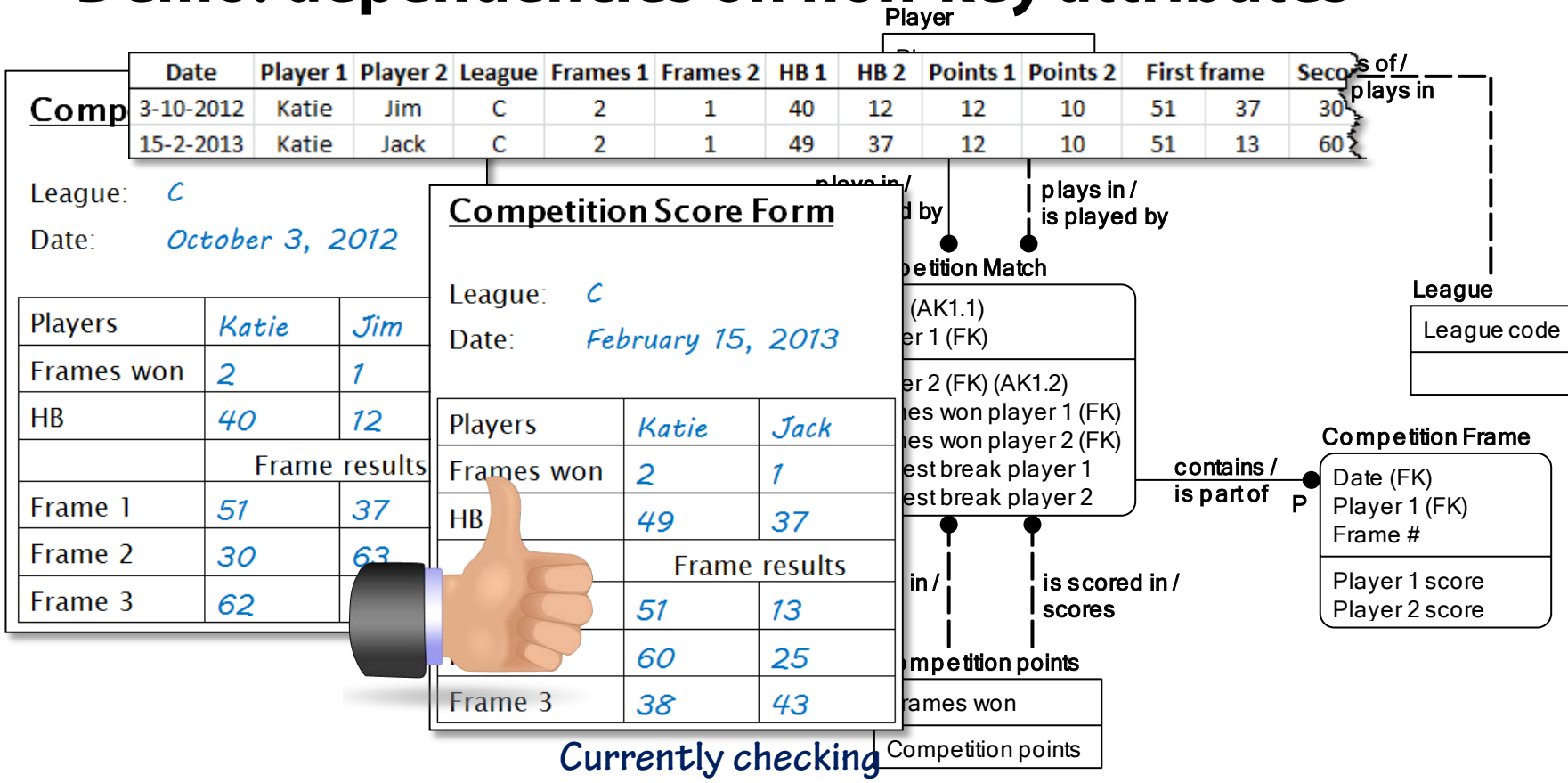
Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2
2012-10-03	Katie	Jim	2	1	40	12
2013-02-15	Katie	Jim	2	1	40	37

Demo: dependencies on non-key attributes



Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2
2012-10-03	Katie	Jim	2	1	40	12
2013-02-15	Katie	Jack	2	1	40	37

Demo: dependencies on non-key attributes

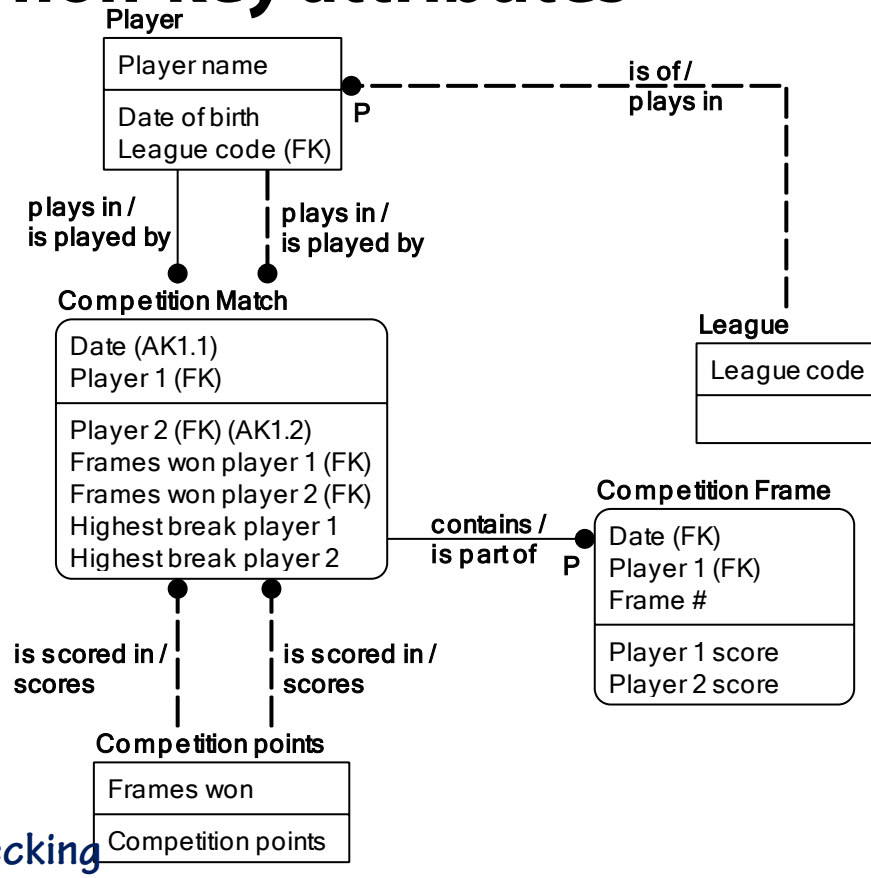


Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2
2012-10-03	Katie	Jim	2	1	40	12
2013-02-15	Katie	Jack	2	1	49	37

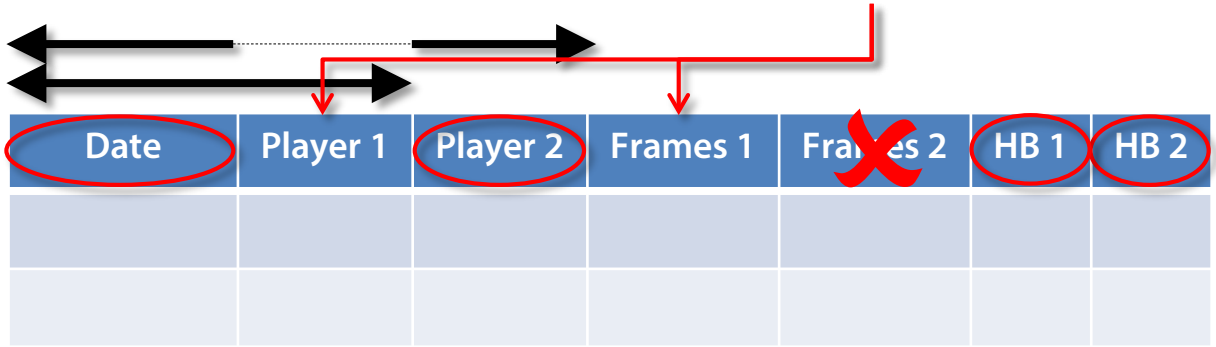
Demo: dependencies on non-key attributes

Dependency found:
- {Player 1, Frames 1} → Frames 2

Actual underlying dependencies:
- {League, Frames 1} → Frames 2
- Player 1 → League



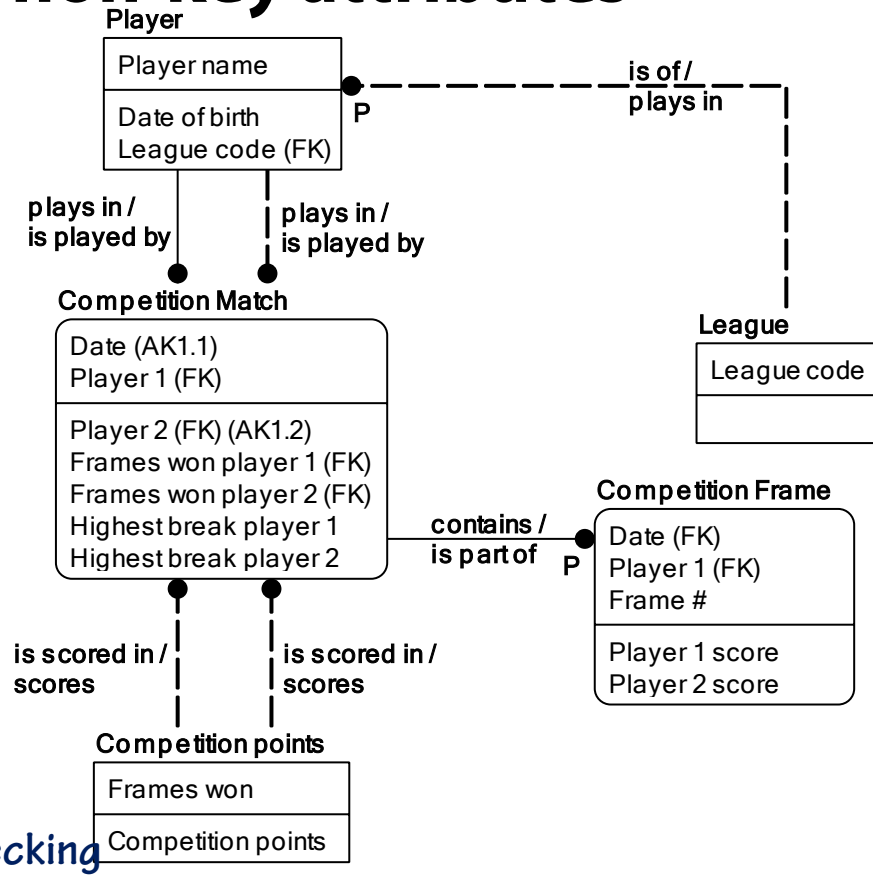
Currently checking



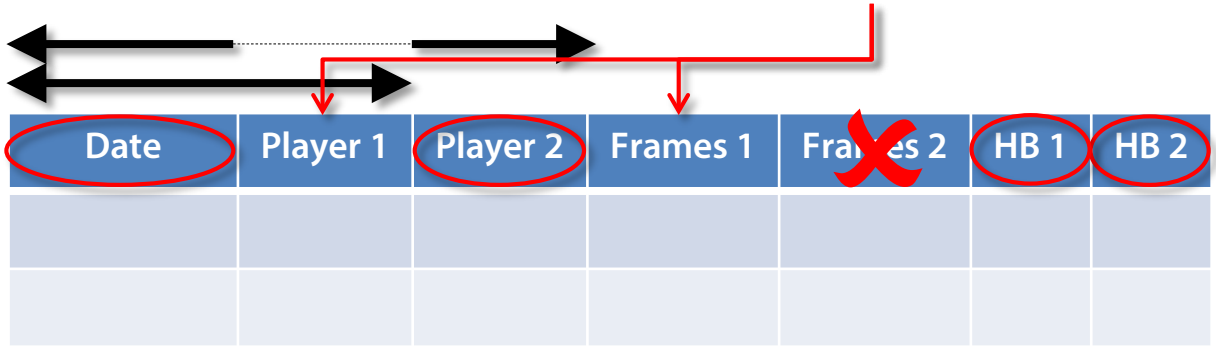
Demo: dependencies on non-key attributes

Dependency found:
- {Player 1, Frames 1} → Frames 2

Functional dependency or derivation rule?



Currently checking

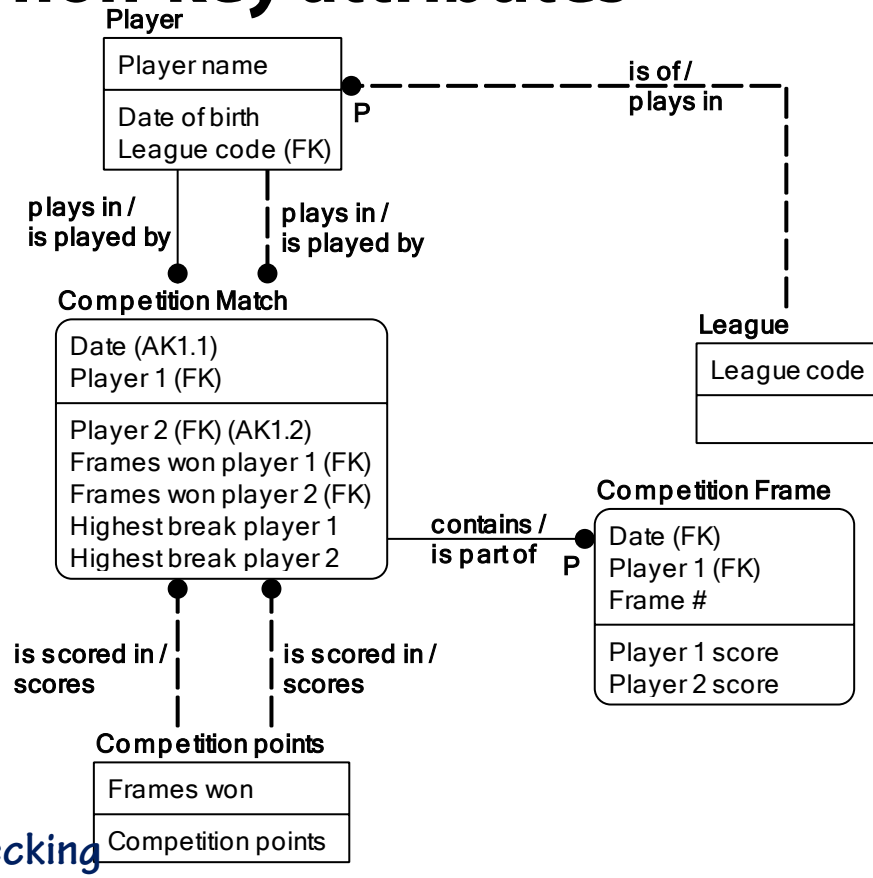


Demo: dependencies on non-key attributes

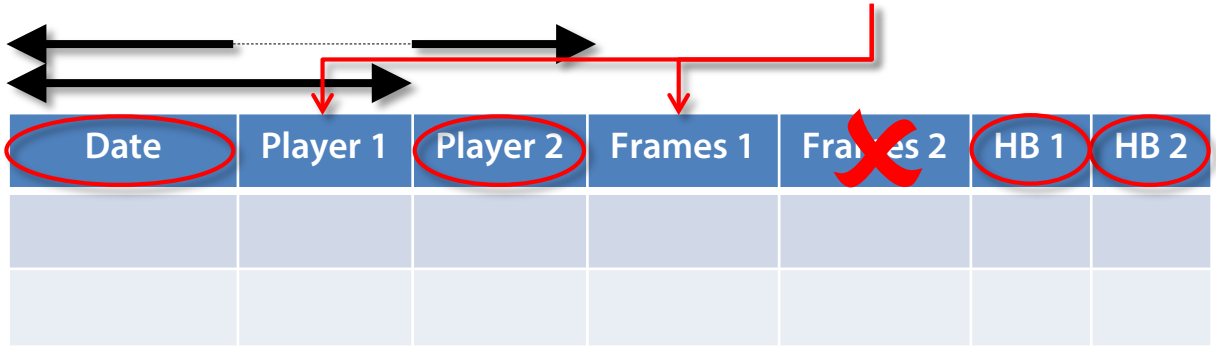
Dependency found:

- {Player 1, Frames 1} → Frames 2

Derivation rule!



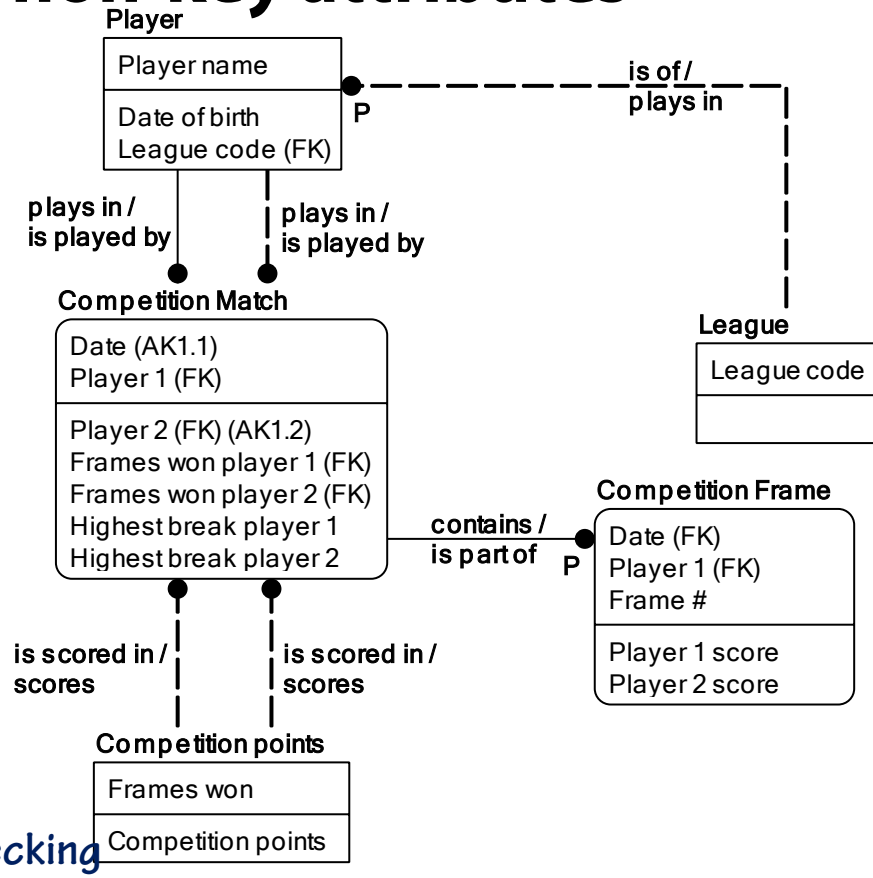
Currently checking



Demo: dependencies on non-key attributes

Dependency found:
- {Player 1, Frames 1} → Frames 2

Derivation rule!

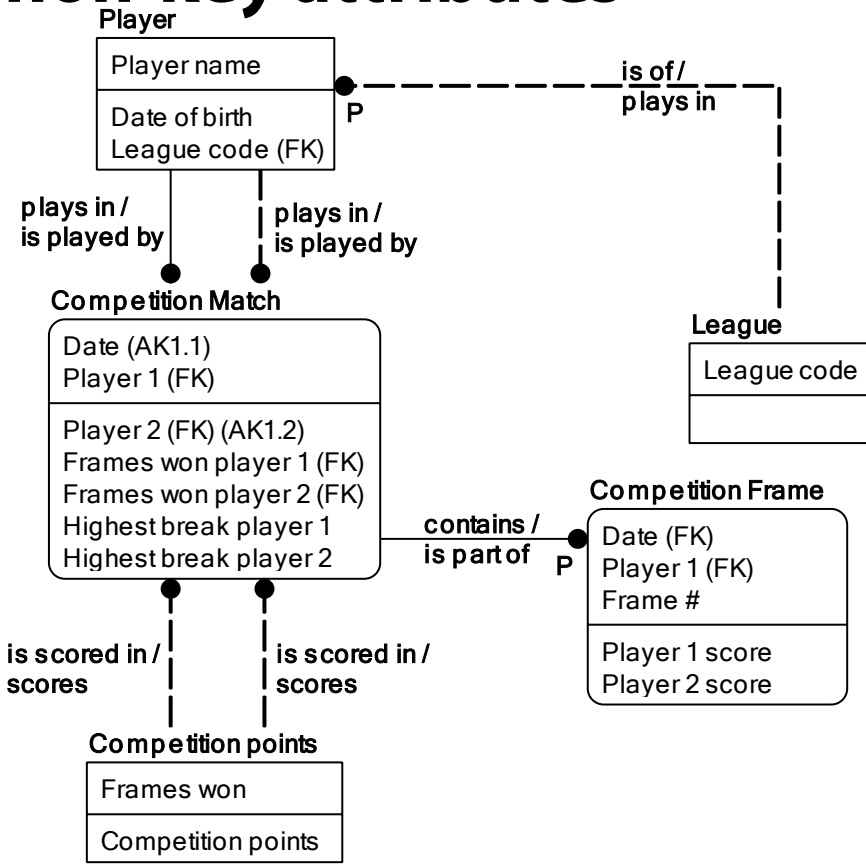


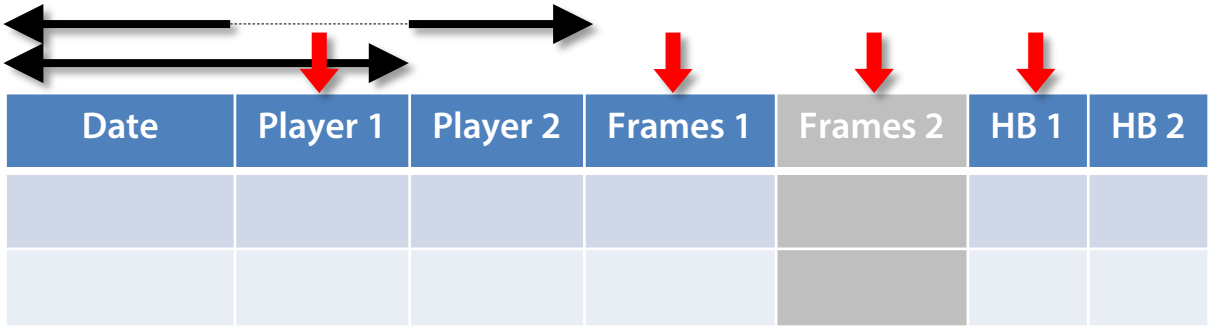
Currently checking

Date	Player 1	Player 2	Frames 1	Frames 2	HB 1	HB 2

Derivation rule:
Frames won player 2
is computed from
Frames won player 1
and
Player 1 → League code
(see specification in design docs)

Demo: dependencies on non-key attributes





Derivation rule:
Frames won player 2
is computed from
Frames won player 1
and
Player 1 → League code
(see specification in design docs)

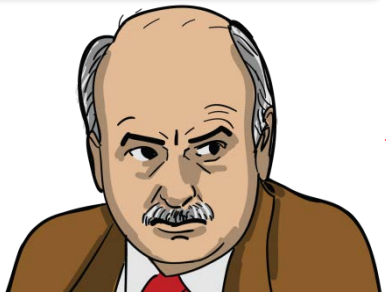
Summary

- **Problems**
  - Redundancy
  - Modification anomalies
  - Fixed by normalization

- **Functional dependencies**

- **Normal forms**
  - First Normal Form
  - Second Normal Form
  - Third Normal Form

The key, the whole key, and nothing but the key; so help me Codd



Dr. E.J. Codd

Illustration: Michael J. Swart

References

- **Further reading:**
  - Database normalization on Wikipedia:
    http://en.wikipedia.org/wiki/Database_normalization
    or
    http://tinyurl.com/Norm-DB

  - P.A. Bernstein's algorithm for synthesis of a Third Normal Form schema:
    http://student.bus.olemiss.edu/files/conlon/others/Others/Bus669_CompInfo/DB_Design_SQL/Bernstein.pdf
    or
    http://tinyurl.com/BernAlgo