On entity resolution in probabilistic data
Ayat, S.N.

Citation for published version (APA):
Ayat, S. N. (2014). On entity resolution in probabilistic data

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Appendix B

The Courses Dataset

This chapter contains the specification of the Courses dataset, which is used for the experiments of Chapter 6.

The Courses dataset contains 17 single-table schemas as follows. Notice that the primary keys are underlined, and $F_i$ is the set of functional dependencies of schema $S_i$, where the functional dependencies that are related to the primary keys are omitted.

$S_1(semester, course, title, instructor, time, room)$
$F_1 = \{course \rightarrow title\}$

$S_2 = (semester, course\#, call\#, name, instructor, days, time, room)$
$F_2 = \{course\# \rightarrow name\}$

$S_3 = (semester, course, sec\#, code\#, room\#, time, days, units, instructor)$
$F_3 = \{code\# \rightarrow \{room\#, time, days, units, instructor\}\}$

$S_4 = (semester, course, crn, time, day, instructor, room, units, title)$
$F_4 = \{course \rightarrow \{units, title\}\}$

$S_5 = (semester, course\#, course, sched\#, days, time, location, instructor)$
$F_5 = \{course\# \rightarrow course\}$
Appendix B. The Courses Dataset

\[ S_6 = \langle \text{term}, \text{coursesubject}, \text{coursenumber}, \text{sectionID}, R, \text{MeetingType}, \text{section}, \text{days}, \text{time}, \text{building}, \text{room}, \text{instructor}, \text{seats-available}, \text{seats-limit}, \text{coursename} \rangle \]

\[ F_6 = \{ \langle \text{coursesubject}, \text{coursenumber} \rangle \rightarrow \text{coursename} \} \]

\[ S_7 = \langle \text{semester}, \text{year}, \text{course-abbr-num}, \text{sec-num}, \text{avl}, \text{enrl-cnt}, \text{type}, \text{course-title}, \text{hr-cr}, \text{time-begin}, \text{time-end}, \text{days}, \text{room}, \text{building}, \text{special-enrollment}, \text{instructor} \rangle \]

\[ F_7 = \{ \langle \text{course-abbr-num} \rangle \rightarrow \{ \text{type}, \text{course-title}, \text{hr-cr} \} \} \]

\[ S_8 = \langle \text{semester}, \text{course#}, \text{ug}, \text{coursename}, \text{teacher}, \text{day}, \text{hour}, \text{place}, \text{comment} \rangle \]

\[ F_8 = \{ \langle \text{course#} \rangle \rightarrow \text{coursename} \} \]

\[ S_9 = \langle \text{term}, \text{course}, \text{subject}, \text{instructor}, \text{time}, \text{room}, \text{units}, \text{grade} \rangle \]

\[ F_9 = \{ \langle \text{course} \rangle \rightarrow \text{subject} \} \]

\[ S_{10} = \langle \text{semester}, \text{course}, \text{title}, \text{units}, \text{lec/sec}, \text{days}, \text{begin}, \text{end}, \text{bldg}, \text{room}, \text{instructors} \rangle \]

\[ F_{10} = \{ \langle \text{course} \rangle \rightarrow \text{title} \} \]

\[ S_{11} = \langle \text{semester}, \text{dept}, \text{course#}, \text{course}, \text{instructor}, \text{day}, \text{time}, \text{room} \rangle \]

\[ F_{11} = \{ \langle \text{dept}, \text{course#} \rangle \rightarrow \text{course} \} \]

\[ S_{12} = \langle \text{semester}, \text{code}, \text{sec}, \text{course}, \text{time}, \text{days}, \text{cr}, \text{room}, \text{instructor} \rangle \]

\[ F_{12} = \{ \langle \text{code} \rangle \rightarrow \text{course} \} \]

\[ S_{13} = \langle \text{semester}, \text{coursecode}, \text{class}, \text{coursename}, \text{remark}, \text{instructor}, \text{time}, \text{building}, \text{roomnumber}, \text{requiredorelective}, \text{numberofstudents} \rangle \]

\[ F_{13} = \{ \langle \text{coursecode} \rangle \rightarrow \text{coursename} \} \]

\[ S_{14} = \langle \text{semester}, \text{course}, \text{sec}, \text{coursename}, \text{days}, \text{time}, \text{cr}, \text{instructor} \rangle \]

\[ F_{14} = \{ \langle \text{course} \rangle \rightarrow \text{coursename} \} \]
$S_{15} = (\text{semester}, \text{course#}, \text{section#}, \text{coursename}, \text{cal#}, \text{days}, \text{times}, \text{room}, \text{status}, \text{max}, \text{now}, \text{instructor}, \text{credits})$

$F_{15} = \{\text{course#} \rightarrow \text{coursename}\}$

$S_{16} = (\text{semester}, \text{coursenumber}, \text{coursetitle}, \text{cr}, \text{days}, \text{time}, \text{bldg}, \text{room}, \text{insr}, \text{ref})$

$F_{16} = \{\text{coursenumber} \rightarrow \text{coursetitle}\}$

$S_{17} = (\text{term}, \text{class#}, \text{title}, \text{location}, \text{day}, \text{times}, \text{instructors}, \text{status})$

$F_{17} = \{\text{class#} \rightarrow \text{title}\}$