CHAPTER IV

FUZZY MULTIVALUED DEPENDENCIES AND INFERENCE RULES

Fuzzy Multivalued Dependencies

The MVDs in classical relational database systems hold in an instance r of R, if for all pairs of tuples $t_1[XYZ] = xy_1z_1$ and $t_2[XYZ] = xy_2z_2$ there exist $t_3[XYZ] = xy_1z_2$ and $t_4[XYZ] = xy_2z_1$ where $X,Y \subseteq R$ and $Z = R - XY$ (Fagin, 1977; Sciore, 1981; Date, 1986; Yang, 1986; Vetter, 1987; Elmasri et.al., 1989). Let $X = A_{i_1}A_{i_2}\ldots A_{i_k}$ and $Y = A_{j_1}A_{j_2}\ldots A_{j_L}$ be subsets of a relation scheme $R(A_1A_2\ldots A_n)$, or simply R. The generalized dependency of MVD : $X \to\to Y$ on fuzzy relational databases called FMVD : $X \sim\to\to Y$ in R is given by:

<u>Definition 4.1.</u> A fuzzy multivalued dependency(FMVD) $X \sim\to\to Y$ with $X, Y \subseteq R$ and $Z = R - XY$, holds in a fuzzy relation r on R, if for all tuples t_1 and t_2 of r such that

$$\mu_{EQ}(t_1[X],t_2[X]) > \max(\mu_{EQ}(t_1[Y],t_2[Y]), \mu_{EQ}(t_1[Z],t_2[Z])), \qquad (4.1)$$

there exists t_3 such that

$$\mu_{EQ}(t_1[Y],t_3[Y]) \geq \mu_{EQ}(t_1[X],t_3[X])) \qquad (4.2)$$

and

$$\mu_{EQ}(t_2[Z],t_3[Z]) \geq \mu_{EQ}(t_2[X],t_3[X])). \qquad (4.3)$$

<u>Remark.</u> The FMVD of the fuzzy relation r on R is assumed to have binary truth values. The MVD in a classical relational database can be viewed as a special case of an FMVD.

Consider the FMVD : $X \sim\!\!\to\!\!\to Y$ and suppose that the relation EQ over dom(A), $A \in XYZ$, satisfies the additional properties that $\mu_{EQ}(a,b) = 0$ for $a \neq b$, $a, b \in dom(A)$. Since for a classical relational database $\mu_r(t) = 1$ for $t \in r$, Definition 4.1 implies that for tuple $t_1[XYZ] = xy_1z_1$ and $t_2[XYZ] = xy_2z_2$ there exists $t_3[XYZ] = xy_1z_2$.

By the symmetry property, it follows this definition that there does not only exist t_3, but also exists t_4 such that

$$\mu_{EQ}(t_1[Z],t_4[Z]) \geq \mu_{EQ}(t_1[X],t_4[X]))$$

and

$$\mu_{EQ}(t_2[Y],t_4[Y]) \geq \mu_{EQ}(t_2[X],t_4[X])).$$

So, FMVD : $X \sim\!\!\to\!\!\to Z$ also holds on R, and it can be inferred that $X \sim\!\!\to\!\!\to Y$ iff $X \sim\!\!\to\!\!\to Z$ for $X, Y, Z \in R$ and $Z = R-XY$.

<u>Theorem 4.1.</u> An FMVD:$X \sim\!\!\to\!\!\to Y$ holds in an instance r of R iff $X \sim\!\!\to\!\!\to Y-X$ holds in r.

<u>Proof.</u> Let $X, Y \subseteq R$ and $Z = R-XY$. <u>Only if part</u>, if $X \sim\!\!\to\!\!\to Y$ holds, then it has to be shown that $X \sim\!\!\to\!\!\to Y-X$ holds. For any two tuples t_1 and t_2, since $Y-X \subseteq Y$, one has

$$\mu_{EQ}(t_1[Y],t_2[Y]) \leq \mu_{EQ}(t_1[Y-X],t_2[Y-X]).$$

Given

$$(1) \qquad \mu_{EQ}(t_1[X],t_2[X]) > \max(\mu_{EQ}(t_1[Y],t_2[Y]), \mu_{EQ}(t_1[Z],t_2[Z])),$$

then there exists t_3 such that

$$(2) \qquad \mu_{EQ}(t_1[X],t_3[X]) \leq \mu_{EQ}(t_1[Y],t_3[Y]),$$

and

$$(3) \qquad \mu_{EQ}(t_2[X],t_3[X]) \leq \mu_{EQ}(t_2[Z],t_3[Z]).$$

It has to be proved that, if

$$(4) \qquad \mu_{EQ}(t_1[X],t_2[X]) > \max(\mu_{EQ}(t_1[Y-X],t_2[Y-X]), \mu_{EQ}(t_1[Z],t_2[Z])),$$

then there exists t_3 such that

$$(5) \qquad \mu_{EQ}(t_1[X],t_3[X]) \leq \mu_{EQ}(t_1[Y-X],t_3[Y-X]),$$

and

(6) $\mu_{EQ}(t_2[X],t_3[X]) \leq \mu_{EQ}(t_2[Z],t_3[Z])$.

It can be seen that whenever (4) is true, (1) is also true. Therefore (4) implies (2) and (3). However (2) and (3) imply (5) and (6), respectively. Hence (4) implies (5) and (6). So, X~->->Y implies X~-->Y-X.

<u>If part</u>, if X~->->Y-X holds, then it has to be shown that X~->->Y holds. From Figure 4.1, one has for any two tuples t' and t''

$$\mu_{EQ}(t'[X],t''[X]) = \min(\mu_{EQ}(t'[X-Y],t''[X-Y]), \mu_{EQ}(t'[X\cap Y],t''[X\cap Y])),$$

and

$$\mu_{EQ}(t'[Y],t''[Y]) = \min(\mu_{EQ}(t'[Y-X],t''[Y-X]), \mu_{EQ}(t'[X\cap Y],t''[X\cap Y])).$$

Never both of these two conditions will hold. Therefor, for any tuples t' and t'', either

(7) $\mu_{EQ}(t_1[X],t_2[X]) \leq \mu_{EQ}(t_1[Y],t_2[Y]) \leq \mu_{EQ}(t_1[Y-X],t_2[Y-X])$,

or

(8) $\mu_{EQ}(t_1[Y],t_2[Y]) \leq \mu_{EQ}(t_1[Y-X],t_2[Y-X]) \leq \mu_{EQ}(t_1[X],t_2[X])$,

depending on whether

$$\mu_{EQ}(t'[X-Y],t''[X-Y]) \leq \mu_{EQ}(t'[Y-X],t''[Y-X]),$$

or

$$\mu_{EQ}(t'[Y-X],t''[Y-X]) \leq \mu_{EQ}(t'[X-Y],t''[X-Y]).$$

By (7) and replacing t' and t'' by t_1 and t_2,respectively, X~->->Y always holds. By (8), whenever (1) is true, (4) is also true. Therefore (1) implies (5) and (6). However, (5) and (6) imply (2) and (3), respectively. Hence (1) implies (2) and (3). Hence, X~->->Y-X implies X~->->Y.

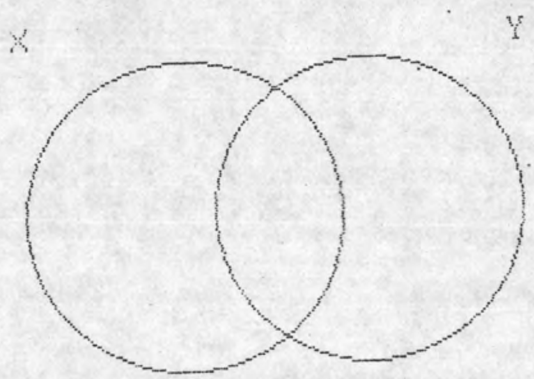

Figure 4.1 Venn diagram for proving Theorem 4.1.

Inference Rules for Fuzzy Multivalued Dependencies

In the following inference rules for FMVDs, X, Y, and Z are subsets of the relation scheme R.

FMVD0(Complementation or Symmetry). Let XYZ = R, X~->->Y iff X~->->Z, where Z = R - XY.

The validity of this rule has been shown in the remark following Definition 4.1.

FMVD1(Reflexivity). If $Y \subseteq X$, then X~->->Y.

Proof. Since Y is a subset of X, and by (3.27), for all tuples t_1 and t_2

$$\mu_{EQ}(t_1[X],t_2[X]) \leq \mu_{EQ}(t_1[Y],t_2[Y]),$$

It follows that (4.1) does not holds, i.e. FMVD:X~->->Y.

<u>FMVD2</u>(Augmentation). If $Z \subseteq W$ and $X \sim\!\!\to\!\!\to Y$, then $XW \sim\!\!\to\!\!\to YZ$.

<u>Proof.</u> By Theorem 4.1, it can be assumed that $X \cap Y = \phi$. For any two tuples t_1 and t_2 such that,

(1) $$\mu_{EQ}(t_1[XW],t_2[XW]) > \max(\mu_{EQ}(t_1[YZ],t_2[YZ]), \mu_{EQ}(t_1[R-XWYZ],t_2[R-XWYZ])),$$

assume that

(2) $$\mu_{EQ}(t_1[X],t_2[X]) > \max(\mu_{EQ}(t_1[Y],t_2[Y]), \mu_{EQ}(t_1[R-XY],t_2[R-XY]))$$

is not true. Then, both

(3) $$\mu_{EQ}(t_1[X],t_2[X]) \leq \mu_{EQ}(t_1[Y],t_2[Y])$$

and

(4) $$\mu_{EQ}(t_1[X],t_2[X]) \leq \mu_{EQ}(t_1[R-XY],t_2[R-XY]))$$

are true.

As $Z \subseteq W$, it follows from (3) that

$$\min(\mu_{EQ}(t_1[X],t_2[X]), \mu_{EQ}(t_1[W],t_2[W])) \leq \min(\mu_{EQ}(t_1[Y],t_2[Y]), \mu_{EQ}(t_1[Z],t_2[Z])),$$

in other words,

(5) $$\mu_{EQ}(t_1[XW],t_2[XW]) \leq \mu_{EQ}(t_1[YZ],t_2[YZ]),$$

and (4) yields

(6) $$\mu_{EQ}(t_1[XW],t_2[XW]) \leq \mu_{EQ}(t_1[Z(R-XY)],t_2[Z(R-XY)]).$$

By the Venn diagram of Figure 4.2, $R-XYZW \subseteq Z(R-XY)$ and (6),

(6') $$\mu_{EQ}(t_1[XW],t_2[XW]) \leq \mu_{EQ}(t_1[R-XYZW],t_2[R-XYZW]).$$

(5) and (6) contradict to initial condition. Hence, if (1) is true, then (2) is true. By (2) and $X \sim\!\!\to\!\!\to Y$, there exists t_3 such that

(7) $$\mu_{EQ}(t_1[X],t_3[X]) \leq \mu_{EQ}(t_1[Y],t_3[Y]),$$

and

(8) $$\mu_{EQ}(t_2[X],t_3[X]) \leq \mu_{EQ}(t_2[R-XY],t_3[R-XY]).$$

By (7), as $Z \subseteq W$, one has

(9) $$\mu_{EQ}(t_1[XW],t_3[XW]) \leq \mu_{EQ}(t_1[YZ],t_3[YZ]).$$

From (8), as $Z \subseteq W$ and $R-XYZW \subseteq Z(R-XY)$, one has

(10) $$\mu_{EQ}(t_2[XW],t_3[XW]) \leq \mu_{EQ}(t_2[R-XYZW],t_3[R-XYZW]).$$

By (1), (9), and (10), $XW \sim\!\!\to\!\!\to YZ$.

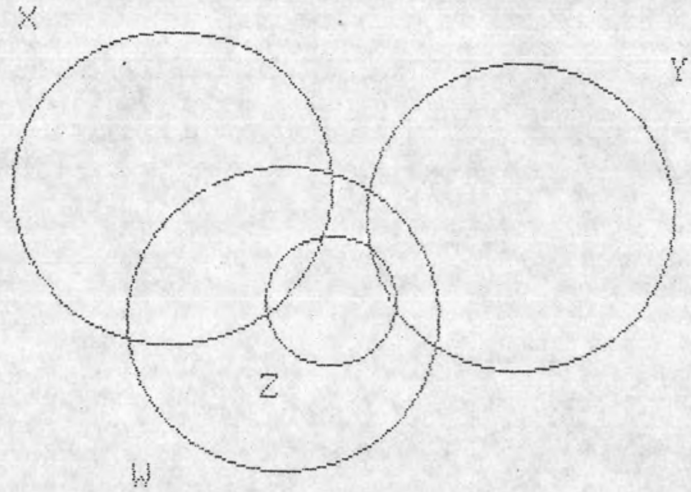

Figure 4.2 Venn diagram for proving FMVD2.

<u>FMVD3</u>(Transitivity). If X~->->Y and Y~->->Z, then X ~->-> Z-Y.

<u>Proof.</u> We proove by contradiction by assuming that the conditions X~->->Y and Y~->->Z are true, but the conclusion X~->->Z-Y is not. By Theorem 4.1, since X~->->Y and Y~->->Z, it can be assumed that $X \cap Y = X \cap Z = \phi$ as shown in Figure 4.3. Given X~->->Y, by FMVD0, one has that X~->->R-XY. Given Y~->->Z, by Theorem 4.1, one has that Y~->->Z-Y, and by FMVD2 and $\phi \subseteq X$, one has that XY~->->Z-Y. If X~->->R-XY holds, then for any two tuples t_1 and t_2 such that

(1) $\mu_{EQ}(t_1[X],t_2[X]) > \max(\mu_{EQ}(t_1[R-XY],t_2[R-XY]), \mu_{EQ}(t_1[Y],t_2[Y]))$,

there exists t_3 such that

(2) $\mu_{EQ}(t_1[X],t_3[X]) \le \mu_{EQ}(t_1[R-XY],t_3[R-XY])$,

and

(3) $\mu_{EQ}(t_2[X],t_3[X]) \le \mu_{EQ}(t_2[Y],t_3[Y])$.

If XY~->->Z-Y holds, then for any two tuples t_4 and t_5 such that

(4) $\mu_{EQ}(t_4[XY],t_5[XY]) > \max(\mu_{EQ}(t_4[Z-Y],t_5[Z-Y]), \mu_{EQ}(t_4[R-XY(Z-Y)],t_5[R-XY(Z-Y)]))$,

there exists t_6 such that

(5) $\mu_{EQ}(t_4[XY],t_6[XY]) \le \mu_{EQ}(t_4[Z-Y],t_6[Z-Y])$,

and

(6) $\mu_{EQ}(t_5[XY],t_6[XY]) \le \mu_{EQ}(t_5[R-XY(Z-Y)],t_6[R-XY(Z-Y)])$.

If X~->->Z-Y does not hold, then

if for any two tuples t_7 and t_8 such that

(7) $\mu_{EQ}(t_7[X],t_8[X]) > \max(\mu_{EQ}(t_7[Z-Y],t_8[Z-Y]),$
$\mu_{EQ}(t_7[R-X(Z-Y)],t_8[R-X(Z-Y)]))$,

then for any t_9

(8) $\mu_{EQ}(t_7[X],t_9[X]) > \mu_{EQ}(t_7[Z-Y],t_9[Z-Y])$,

or

(9) $\mu_{EQ}(t_8[X],t_9[X]) > \mu_{EQ}(t_8[R-X(Z-Y)],t_9[R-X(Z-Y)])$.

Assume there exist $t_1 = t_4$ and $t_2 = t_5$ such that (1) and (4) are true. As $Y \subseteq XY$, from (1) and (4) one has

$$\mu_{EQ}(t_1[X],t_2[X]) > \max(\mu_{EQ}(t_1[R-XY],t_2[R-XY]), \max(\mu_{EQ}(t_1[Z-Y],t_2[Z-Y]), \mu_{EQ}(t_1[R-XY(Z-Y)],t_2[R-XY(Z-Y)])));$$

since $R-XY(Z-Y) \subseteq R-XY$, one has

$$\mu_{EQ}(t_1[X],t_2[X]) > \max(\mu_{EQ}(t_1[Z-Y],t_2[Z-Y]), \mu_{EQ}(t_1[R-XY(Z-Y)],t_2[R-XY(Z-Y)])),$$

and as $R-XY(Z-Y) \subseteq R-X(Z-Y)$, one has

$$\mu_{EQ}(t_1[X],t_2[X]) > \max(\mu_{EQ}(t_1[Z-Y],t_2[Z-Y]), \mu_{EQ}(t_1[R-X(Z-Y)],t_2[R-X(Z-Y)]));$$

hence, whenever (1) and (4) are true, (7) is also true (replacing t_7 and t_8 by t_1 and t_2, respectively). Now, replace t_7 and t_8 in (8) and (9) by t_1 and t_2, respectively, and as $Z-Y \subseteq R-XY$, by (8), one has

$$\mu_{EQ}(t_1[X],t_9[X]) > \mu_{EQ}(t_1[R-XY],t_9[R-XY]),$$

which contradicts with (2). Therefore, (8) cannot hold. Moreover, if (9) is true, as $Y \subseteq R-X(Z-Y)$, then for any t_9

$$\mu_{EQ}(t_2[X],t_9[X]) > \mu_{EQ}(t_2[Y],t_9[Y]),$$

which contradicts with (3). Therefore, (9) cannot hold. Therefore, it is not true that X~->->Z-Y does not hold when X~->->Y and Y~->->Z are true. In other words, we have proved that X~->->Y and Y~->->Z imply X~->->Z-Y.

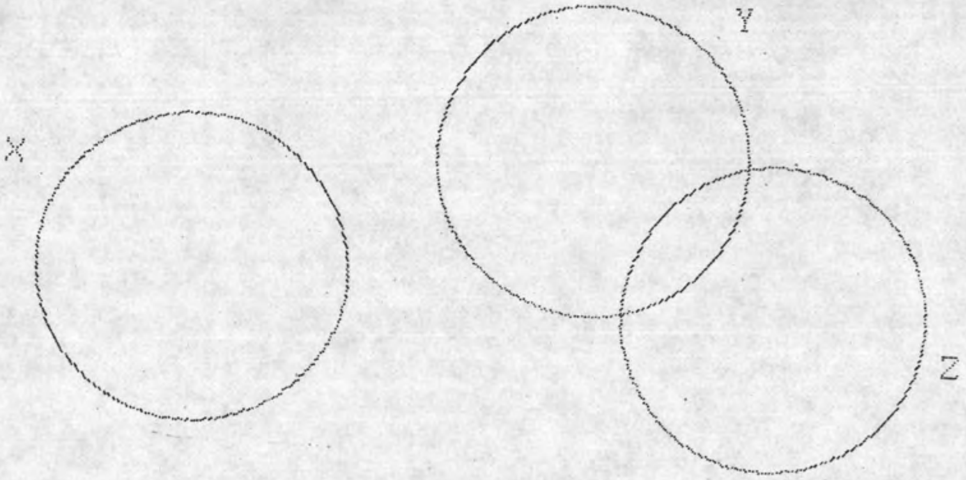

Figure 4.3 Venn diagram for proving FMVD3.

The following inference rules follow from the above rules

FMVD4(Pseudo-Transitivity). If X~->->Y and YW~->->Z, then XW ~->-> Z-YW.

Proof.

(1) X~->->Y (given FMVD),

(2) XW~->->YW (by applying FMVD2 to (1)),

(3) YW~->->Z (given FMVD),

(4) XW~->->Z-YW(by applying FMVD3 on (2) and (3)).

FMVD5(Union or Additivity). If X~->->Y and X~->->Z, then X ~->->YZ.

Proof. By Theorem 4.1, it can be assume that $X \cap Y = X \cap Z = \phi$,

(1) X~->->Z (given FMVD),

(2) X~->->XZ (by applying FMVD2),

(3) X~->->Y (given FMVD),

(4) XZ~->->YZ (by applying FMVD2),

(5) XZ~->->R-XYZ (by applying FMVD0),

(6) X~->->R-XYZ (by applying FMVD3 to (2) and (5)),

(7) X~->->R-X(R-XYZ) (by applying FMVD0),

(8) X~->->YZ (YZ=R-X(R-XYZ)).

FMVD6(Decomposition and projectivity). If X~->->Y and X~->->Z, then X~->->Y∩Z, X~->->Y-Z and X~->->Z-Y.

Proof. By Theorem 4.1, it can be assumed that X∩Y=X∩Z=ϕ as shown in Figure 4.3,

(1) X~->->Y (given FMVD),

(2) X~->->Z (given FMVD),

(3) X~->->R-XY (by applying FMVD0 on X~->->Y),

(4) X~->->Z(R-XY) (by applying FMVD5 on (2) and (3)),

(5) X~->->Y-Z (by applying FMVD0 on (4)),

(6) X~->->R-XZ (by applying FMVD0 on (2)),

(7) X~->->Y(R-XZ) (by applying FMVD5 on (1) and (6)),

(8) X~->->Z-Y (by applying FMVD0 on (7)),

(9) X~->->(R-XZ)(R-XY) (by applying FMVD5 on (3) and (6)),

(10) X~->->Y∩Z (by applying FMVD0 on (9)).

Inference Rules for Functional and Multivalued Dependencies

Let W, X, Y and Z be subsets of R.

FFM1(Replication). If X~->Y, then X~->->Y.

Proof. Since X~->Y, then for any two tuples t_1 and t_2

$$\mu_{EQ}(t_1[X],t_2[X]) \leq \mu_{EQ}(t_1[Y],t_2[Y]).$$

Hence, (4.1) does not hold, i.e. FMVD:X~->->Y holds.

FFM2(Coalescence). If X~->->Y and Z~->W where W $\subseteq$ Y and Y∩Z = ϕ, then X~->W.

Proof. Since X~->->Y and W$\subseteq$Y, by Theorem 4.1, assume that X∩Y=Y∩Z=ϕ, as shown in Figure 4.4. If X~->W does not holds and Z~->W holds, then for any two tuples t_1 and t_2

(1) $$\mu_{EQ}(t_1[X]),t_2[X]) > \mu_{EQ}(t_1[W],t_2[W]),$$

and

(2) $$\mu_{EQ}(t_1[W]),t_2[W]) \geq \mu_{EQ}(t_1[Z],t_2[Z]).$$

By (1), as $W \subseteq Y$,

(3) $$\mu_{EQ}(t_1[X]),t_2[X]) > \mu_{EQ}(t_1[Y],t_2[Y]).$$

By (1), (2), and $Z \subseteq (X \cap Z)(R-XY)$,

$$\mu_{EQ}(t_1[X]),t_2[X]) > \mu_{EQ}(t_1[(X \cap Z)(R-XY)],t_2[(X \cap Z)(R-XY)]),$$

i.e.

$$\mu_{EQ}(t_1[X]),t_2[X]) > \min(\mu_{EQ}(t_1[X \cap Z],t_2[X \cap Z),$$
$$\mu_{EQ}(t_1[R-XY],t_2[R-XY)),$$

but $X \cap Z \subseteq X$,

$$\mu_{EQ}(t_1[X]),t_2[X]) \leq \mu_{EQ}(t_1[X \cap Z],t_2[X \cap Z],$$

then always

(4) $$\mu_{EQ}(t_1[X]),t_2[X]) > \mu_{EQ}(t_1[(R-XY)],t_2[(R-XY)].$$

It been shown that any t_1 and t_2 must satisfy (3) and (4), then for any t_1 and t_2 such that

$$\mu_{EQ}(t_1[X]),t_2[X]) > \max(\mu_{EQ}(t_1[Y],t_2[Y]),$$
$$\mu_{EQ}(t_1[R-XY],t_2[R-XY])),$$

there does not exist t_3 such that

$$\mu_{EQ}(t_1[X]),t_3[X]) \leq \mu_{EQ}(t_1[Y],t_3[Y]).$$

Hence, $X \sim\!\!->\!->Y$ does not hold. So, $X \sim\!\!->\!->Y$ and $Z \sim\!\!->W$, where $W \subseteq Y$ and $Y \cap Z = \phi$ imply $X \sim\!\!->W$.

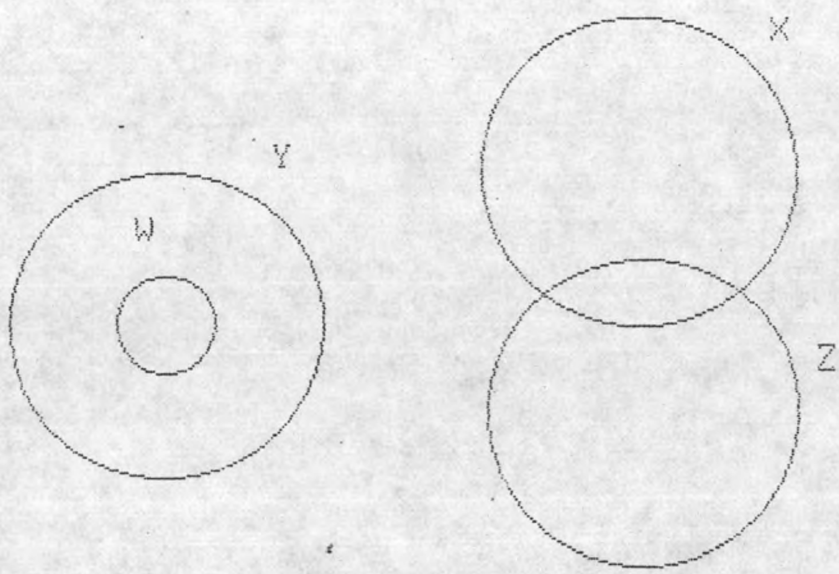

Figure 4.4 Venn diagram for proving FFM2.