Exercises on Sequential pattern mining

Ex. 1

< {A}	{A,B,F}	{B,F}	{E,F}	{G}	{A}	{A,B}	{E} >
t=0	t=1	t=2	t=3	t=4	t=5	t=6	t=7

show all the occurrences (there can be more than one or none, in general) of each of the following subsequences in the input sequence above. Repeat the exercise twice: the first time considering no temporal constraints (left column): the second time considering max-gap = 2 (i.e. gap <= 2, right column). Each occurrence should be represented by its corresponding list of time stamps, e.g..: <0,2,3> = <t=0, t=2, t=3>.

Solutions are highlighted in yellow:

	Occurrences	Occurrences with max-gap = 2
ex.: <{B}{E}>	<1,3> <1,7> <2,3><2,7><6,7>	<1,3> <2,3><6,7>
w_1 = <{A} {B} {E} >	<0,1,3><0,1,7><0,2,3> <0,2,7><0,6,7> <1,2,3> <1,2,7><1,6,7> <5,6,7>	<0,1,3><0,2,3> <1,2,3> <5,6,7>
w_2 = <{B}{G}{A}>	<1,4,5><1,4,6> <2,4,5> <2,4,6>	<2,4,5> <2,4,6>
w_3 = <{A,B} {E}>	<1,3> <1,7> <6,7>	<1,3> <6,7>

Ex. 2

Given the input sequences listed in the table below (column 1), show for each of them **all the occurrences** of subsequences {A} → {D} and {A} → {C,D}, and finally write its total support. Repeat the exercise twice: the first time **considering no temporal constraints** (columns 2 and 4); the second time **considering min-gap = 1** (i.e. gap > 1) (columns 3 and 5). Each occurrence should be represented by its corresponding list of time stamps, e.g..: <0,2,3> = <t=0, t=2, t=3>.

column 1	column 2	column 3	column 4	column 5
	{A} → {D}		{B} → {C,D}	
	No constraints	min-gap = 1	No constraints	min-gap = 1
< {A,B,F} {C} {C,D,F} {E} {C,D} > t=0 t=1 t=2 t=3 t=4	<0,2>, <0,4>	<0,2> <0,4>	<0,2>, <0,4>	<0,2>, <0,4>
< {A,B} {C} {A,B} {C,D} > t=0 t=1 t=2 t=3	<0,3> <2,3>	<0,3>	<0,3>, <2,3>	<0,3>
< {F} {A,B,F} {A,B,C,D} {D} {E} {C} > t=0 t=1 t=2 t=3 t=4 t=5	<1,2> <1,3> <2,3>	<1,3>	<1,2>	none
< {A,F} {B,C} {A,B} {E} {D} > t=0 t=1 t=2 t=3 t=4	<0,4> <2,4>	<0,4> <2,4>	none	none
< {A,B,F} {A,C} {A,B,D} {C} {C,D} > t=0 t=1 t=2 t=3 t=4	<0,2> <0,4> <1,2>, <1,4> <2,4>	<0,2> <0,4> <1,4> <2,4>	<0,4> <2,4>	<0,4> <2,4>
Total support:	5 (100%)	5 (100%)	4 (80%)	3 (60%)