## SAMPLE SET

(Aquellos textos en los que tan solo se menciona el identificador y no aparecen, bien subrayados, bien con anotaciones, son textos procesados correctamente.)

**SUBTRACK 1**

S0004-06142005001000011-1.txt

S0004-06142005001000011-3.txt

S0004-06142005001000016-1.txt

S0004-06142006000600014-1.txt //Solucionado null pointer debido a nada detectado por el analizador sintactico (lambda)

S0004-06142006000600015-1.txt

S0004-06142006000200014-1.txt

S0004-06142006000300015-1.txt

S0004-06142007000500011-1.txt //VSG de más. No se puede detectar

S0004-06142006000700013-1.txt

S0004-06142006000700014-1.txt //Saca TAC de mas

S0004-06142006000900015-1.txt

S0004-06142007000600012-1.txt

S0004-06142007000600016-2.txt

S0004-06142007000700014-1.txt

S0004-06142007000900013-1.txt

## TRAINING SET

**SUBTRACK 1**

S0004-06142005000900013-1.txt

S0212-71992007000100006-1.txt

S1130-05582008000400008-3.txt

S0004-06142005000900015-1.txt

S0212-71992007000200007-1.txt

S1130-05582008000500006-1.txt

S0004-06142005000900016-1.txt

S0212-71992007000400005-1.txt

S1130-05582008000600006-1.txt

S0004-06142005001000011-1.txt

S0212-71992007001000007-1.txt

S1130-05582009000400005-1.txt

S1130-05582009000600006-1.txt

S0004-06142005001000016-1.txt

S0213-12852003000600002-1.txt

S1130-05582010000200003-1.txt

S0004-06142006000200013-1.txt

S0213-12852004000500002-1.txt

S1130-05582010000200004-1.txt

S0004-06142006000200014-1.txt

S0213-12852016000400002-1.txt

S1130-05582012000300005-1.txt

S0004-06142006000300015-1.txt

S0213-12852016000500002-1.txt

S1130-05582013000200007-1.txt

S0004-06142006000600014-1.txt

S0365-66912004000800009-1.txt

S1130-05582013000400005-1.txt //Detecta dos formas para TAC una de ellas una frase completa, la otra la forma correcta

S0004-06142006000600015-1.txt

S0365-66912004001000008-1.txt

S1130-05582014000300006-1.txt

S0004-06142006000700008-1.txt

S0365-66912005000600011-1.txt

S1130-05582014000400006-1.txt

S0004-06142006000700013-1.txt

S0365-66912005001000011-1.txt

S1130-05582015000200006-1.txt

S0004-06142006000700014-1.txt //Mala detección de TAC, obtiene frase completa

S0365-66912005001100009-1.txt

S1130-05582017000100037-1.txt

S0004-06142006000900006-1.txt

S0365-66912005001200010-1.txt

S1130-05582017000100044-2.txt

S0004-06142006000900015-1.txt

S0365-66912006000100008-1.txt

S1130-05582017000200099-1.txt

S0004-06142007000300012-1.txt

S0365-66912006000200012-1.txt

S1130-05582017000200117-1.txt

S0004-06142007000500011-1.txt //Forma VSG errónea, saca frase

S0365-66912006000500008-1.txt

S1130-05582017000200119-1.txt

S0004-06142007000600012-1.txt

S0365-66912006000600011-1.txt

S1130-05582017000300150-3.txt

S0004-06142007000600016-2.txt

S0365-66912006000900010-1.txt

S1130-14732005000600008-1.txt

S0004-06142007000700014-1.txt

S0365-66912006001100006-4.txt

S1130-14732006000500005-1.txt

S0004-06142007000900010-1.txt

S0365-66912006001200008-1.txt

S1130-63432013000100011-1.txt //Solucionar chi fallo de caracter

S0004-06142007000900013-1.txt

S0365-66912006001200009-1.txt

S1130-63432013000300013-1.txt

S0004-06142008000400011-2.txt

S0365-66912007000200011-1.txt

S1130-63432013000600017-1.txt

S0004-06142008000400013-1.txt

S0365-66912007000800013-1.txt

S1130-63432014000100012-1.txt

S0004-06142008000500014-1.txt

S0365-66912007001200011-1.txt

S1130-63432014000200011-1.txt

S0004-06142008000500015-1.txt

S0365-66912009000300009-1.txt

S1130-63432014000300013-4.txt

S0004-06142008000600013-1.txt

S0365-66912009000300010-1.txt

S1130-63432014000400015-1.txt

S0004-06142008000600014-1.txt

S0365-66912009000800005-1.txt

S1130-63432014000500008-1.txt

S0004-06142008000700011-1.txt

S0365-66912009001100005-1.txt

S1130-63432016000100009-1.txt

S0004-06142008000700015-2.txt

S0365-66912009001100005-3.txt

S1130-63432016000500012-1.txt

S0004-06142009000100010-3.txt

S0365-66912009001100006-1.txt

S1131-57682001000100007-2.txt

S0004-06142009000100012-1.txt

S0365-66912010001200005-1.txt

S1131-57682003000300008-1.txt // SOLUCIONADO: error si ac long 2 en check from start TB mal detectado.

S0004-06142009000100013-1.txt

S0365-66912011000400005-2.txt

S1131-57682003000400014-1.txt

S0004-06142009000300014-1.txt

S0365-66912011000900005-1.txt

S1134-80462004000600007-1.txt

S0004-06142009000500011-1.txt

S0365-66912011001000003-1.txt

S1134-80462006000600005-1.txt

S0004-06142009000800011-1.txt

S0365-66912012000200004-1.txt

S1134-80462008000200008-1.txt

S0004-06142010000100015-1.txt

S0365-66912012000300005-1.txt

S1134-80462008000800003-1.txt

S0210-48062003000400012-1.txt

S0365-66912012000500005-1.txt

S1134-80462009000100005-16.txt //Solucionado: algunos caracteres no se saben cuáles son exactamente como micro, por lo que no se puede incluir su codificación en Unicode en el léxico

S0210-48062003000900013-1.txt

S0365-66912012000500005-2.txt

S1134-80462009000500006-1.txt

S0210-48062004000500008-1.txt

S0365-66912012000800005-1.txt

S1134-80462014000300004-4.txt

S0210-48062004000900010-1.txt

S0376-78922007000400008-1.txt

S1134-80462015000100006-1.txt

S0210-48062005000700010-1.txt

S0376-78922008000400008-1.txt

S1134-80462015000200005-1.txt

S0210-48062005000700015-1.txt

S0376-78922009000100011-1.txt

S1135-76062005000100002-1.txt

S0210-48062006000100012-1.txt

S0376-78922009000100014-1.txt

S1135-76062009000300004-1.txt

S0210-48062006000200018-1.txt

S0376-78922009000200008-1.txt

S1135-76062009000400004-1.txt

S0210-48062006000400012-1.txt

S0376-78922009000300010-1.txt

S1135-76062010000200004-1.txt

S0210-48062006000700010-1.txt

S0376-78922009000400002-1.txt

S1135-76062010000200006-1.txt

S0210-48062006000900012-1.txt

S0376-78922009000400010-1.txt

S1135-76062011000100006-1.txt

S0210-48062007000300014-1.txt

S0376-78922011000500005-1.txt

S1137-66272004000500009-1.txt

S0210-48062007000400012-1.txt

S0376-78922011000500012-1.txt

S1137-66272006000300009-1.txt

S0210-48062007000400013-1.txt

S0376-78922012000100001-1.txt

S1137-66272007000300014-1.txt

S0210-48062007000700016-1.txt

S0376-78922012000200008-1.txt

S1137-66272007000500016-1.txt

S0210-48062007000700017-1.txt //Error con BCG

S0376-78922012000300005-1.txt

S1137-66272009000500017-1.txt

S0210-48062007001000004-3.txt

S0376-78922012000300010-1.txt

S1137-66272011000100012-1.txt

S0210-48062007001000004-4.txt

S0376-78922014000200012-1.txt

S1137-66272011000100013-3.txt

S0210-48062007001000016-1.txt

S0376-78922014000300013-1.txt

S1137-66272011000300023-1.txt

S0210-48062009000200019-1.txt

S0376-78922016000200011-1.txt

S1137-66272012000100017-1.txt

S0210-48062009000300017-1.txt

S0376-78922016000300010-1.txt

S1137-66272012000300023-1.txt

S0210-48062009000400017-1.txt

S0376-78922017000100008-1.txt

S1137-66272012000300024-1.txt

S0210-48062009000900015-1.txt

S0378-48352004000300007-1.txt

S1137-66272013000100023-1.txt

S0210-48062009000900015-3.txt

S0378-48352004000900006-1.txt

S1137-66272013000200023-1.txt

S0210-48062010000100019-3.txt

S0378-48352005000400004-1.txt

S1137-66272014000100019-1.txt

S0210-48062010000100019-4.txt

S0378-48352006000300005-1.txt

S1137-66272015000200019-2.txt

S0210-56912006000600007-1.txt

S0378-48352006000900006-1.txt

S1137-66272015000300014-1.txt

S0210-56912006000800008-1.txt

S0465-546X2009000300008-1.txt

S1137-66272016000200010-1.txt

S0210-56912007000200007-3.txt

S0465-546X2010000200006-1.txt

S1137-66272016000200015-1.txt

S0210-56912007000700008-1.txt

S0465-546X2014000400012-1.txt

S1137-66272016000300012-1.txt

S0210-56912007000900007-1.txt

S1130-01082004001200010-1.txt

S1138-123X2004000100006-2.txt

S0210-56912008000400007-3.txt

S1130-01082005000500013-1.txt

S1139-13752009000200010-1.txt

S0210-56912008000400007-4.txt

S1130-01082005001100016-1.txt

S1139-13752009000200010-2.txt

S0210-56912010000200009-1.txt //Fallo character no reconocible en utf8 (micro)

S1130-01082006000100014-1.txt

S1139-76322009000100006-1.txt

S0210-56912011000700007-1.txt

S1130-01082007000400010-1.txt

S1139-76322009000700014-2.txt

S0210-56912012000900010-1.txt

S1130-01082007000400015-1.txt

S1139-76322010000600010-1.txt

S0211-69952009000500014-1.txt

S1130-01082007000400016-1.txt

S1139-76322011000100011-1.txt

S0211-69952009000600023-1.txt

S1130-01082007000700011-1.txt

S1139-76322012000300006-2.txt

S0211-69952011000500011-1.txt

S1130-01082007000700011-2.txt

S1139-76322012000400010-1.txt

S0211-69952011000500011-3.txt

S1130-01082007001000017-1.txt

S1139-76322013000400008-1.txt

S0211-69952011000500013-1.txt

S1130-01082007001100010-1.txt

S1139-76322014000500014-1.txt

S0211-69952012000700031-1.txt

S1130-01082007001100010-2.txt

S1139-76322015000100010-1.txt

S0211-69952013000200025-1.txt

S1130-01082007001100012-1.txt

S1139-76322015000100012-4.txt

S0211-69952014000200012-1.txt

S1130-01082007001200012-1.txt

S1139-76322015000300015-1.txt

S0211-69952014000200018-1.txt

S1130-01082008000400013-1.txt

S1139-76322015000400006-1.txt

S0211-69952014000400019-1.txt

S1130-01082008000400017-1.txt

S1139-76322015000500009-1.txt

S0211-69952015000200015-1.txt

S1130-01082008000500008-1.txt

S1139-76322016000300007-1.txt

S0211-69952015000300011-1.txt

S1130-01082008000800009-1.txt

S1139-76322016000400012-1.txt

S0211-69952016000300015-1.txt

S1130-01082008000800010-1.txt

S1139-76322017000200007-1.txt //Revisar salida

S0212-16112007000600015-1.txt

S1130-01082008000900014-1.txt

S1698-44472004000200009-1.txt

S0212-16112007000800011-1.txt

S1130-01082008001000008-1.txt

S1698-44472004000400012-1.txt

S0212-16112009000300015-1.txt

S1130-01082008001100009-1.txt

S1698-44472005000400012-1.txt

S0212-16112009000600017-1.txt

S1130-01082008001200016-2.txt

S1698-69462006000100006-1.txt

S0212-16112010000100017-1.txt

S1130-01082009000200012-1.txt

S1698-69462006000300015-1.txt

S0212-16112011000600041-1.txt

S1130-01082009000300015-1.txt

S1699-695X2014000200007-1.txt

S0212-16112012000600045-1.txt

S1130-01082009000400009-1.txt

S1699-695X2015000100013-1.txt

S0212-16112013000300055-1.txt

S1130-01082009000400011-1.txt

S1699-695X2015000200010-1.txt

S0212-16112013000400053-1.txt

S1130-01082009000400014-1.txt

S1699-695X2015000300010-1.txt //Mirar acronimos con m

S0212-16112014000200032-3.txt

S1130-01082009000400015-1.txt

S1699-695X2016000200009-1.txt

S0212-71992003000500006-1.txt

S1130-01082009000700016-1.txt // Error PCR reactiva. Debido a que es proteína C reactiva, si se incluye la mayuscula puede dar lugar a ignorar acronimos en la subtarea 2

S1887-85712011000200007-1.txt

S0212-71992005000400007-1.txt

S1130-01082009001000013-1.txt

S1887-85712012000400006-1.txt

S0212-71992005000400009-1.txt

S1130-01082010000600012-1.txt

S1887-85712014000100005-1.txt

S0212-71992005001000009-1.txt

S1130-01082011000500014-1.txt

S1887-85712015000100005-1.txt

S0212-71992006000100008-1.txt

S1130-05582003000600004-3.txt

S1887-85712015000200005-1.txt

S0212-71992006000600008-1.txt

S1130-05582007000500003-1.txt

S1888-75462016000400180-1.txt

S0212-71992006000700009-1.txt

S1130-05582007000600003-1.txt

S1889-836X2015000100003-1.txt

S0212-71992006000800007-1.txt

S1130-05582008000200006-1.txt

S1889-836X2015000200005-1.txt

S0212-71992006000800008-1.txt

S1130-05582008000400007-2.txt

S2340-98942015000100005-1.txt

Mirar fallo acronimos TAP trasabdomial percutaneo

**TESTING SET**

S0004-06142005000900014-1.txt

S0004-06142006000500002-4.txt

S0004-06142006000500012-1.txt

S0004-06142006000600012-1.txt

S0004-06142007000600011-1.txt

S0004-06142007000700013-1.txt

S0004-06142007000900012-1.txt

S0004-06142007000900014-1.txt

S0004-06142008000300015-1.txt

S0004-06142009000100009-1.txt

S0004-06142009000200008-1.txt

S0004-06142009000200013-1.txt

S0004-06142009000400011-1.txt

S0004-06142009000500009-1.txt

S0004-06142010000500014-1.txt

S0210-48062003000800013-1.txt

S0210-48062003000900009-1.txt

S0210-48062004000200011-1.txt

S0210-48062005000300016-2.txt

S0210-48062005000600012-1.txt

S0210-48062005000800014-1.txt

S0210-48062006000300014-1.txt

S0210-48062006000300015-1.txt

S0210-48062006000900014-1.txt

S0210-48062007000100004-1.txt

S0210-48062007000100011-1.txt

S0210-48062007000500015-1.txt

S0210-48062007000700014-1.txt

S0210-48062007000700015-1.txt

S0210-48062009000200017-1.txt

S0210-48062009000300013-4.txt

S0210-48062009000400020-1.txt

S0210-56912008000100008-3.txt

S0210-56912008000200007-2.txt

S0210-56912011000500009-3.txt

S0211-57352013000300011-1.txt

S0211-57352013000300012-1.txt

S0211-69952010000400013-1.txt

S0211-69952011000400013-1.txt

S0211-69952012000500025-1.txt

S0211-69952012000700028-1.txt

S0211-69952013000200018-1.txt

S0211-69952013000200019-1.txt

S0211-69952013000500019-1.txt

S0211-69952013000500035-1.txt

S0211-69952014000100016-1.txt

S0211-69952015000100013-1.txt

S0211-69952016000200176-1.txt

S0211-69952016000600552-1.txt

S0211-69952017000200225-2.txt

S0212-16112004000400007-1.txt

S0212-16112005000600011-1.txt

S0212-16112007000800012-1.txt

S0212-16112011000300031-1.txt

S0212-16112012000400060-1.txt

S0212-16112014000300032-1.txt

S0212-16112014000600030-2.txt

S0212-71992000001200010-1.txt

S0212-71992004000300009-1.txt

S0212-71992004001100006-1.txt

S0212-71992004001100007-1.txt

S0212-71992005000200008-1.txt

S0212-71992005000500009-1.txt

S0212-71992005000600008-1.txt

S0212-71992005000700008-1.txt

S0212-71992005001200010-1.txt

S0212-71992006000600007-1.txt

S0212-71992006000900007-1.txt

S0212-71992007000300007-1.txt

S0212-71992007000700008-1.txt

S0212-71992007001000008-1.txt

S0213-12852005000200004-1.txt

S0365-66912004000600008-1.txt

S0365-66912004000900009-1.txt

S0365-66912004001200011-1.txt

S0365-66912005001100010-2.txt

S0365-66912005001200008-1.txt

S0365-66912006000300009-1.txt

S0365-66912006000600010-2.txt

S0365-66912006000700011-1.txt

S0365-66912006000900011-1.txt

S0365-66912006001000011-1.txt

S0365-66912006001100007-1.txt

S0365-66912006001200010-1.txt

S0365-66912007000200009-1.txt

S0365-66912007000700011-1.txt

S0365-66912007000700012-1.txt

S0365-66912007000900009-1.txt

S0365-66912007001000008-1.txt

S0365-66912007001100010-1.txt

S0365-66912007001200009-1.txt

S0365-66912007001200012-1.txt

S0365-66912007001200013-1.txt

S0365-66912008000700008-1.txt

S0365-66912008001000010-1.txt

S0365-66912010000300005-1.txt

S0365-66912010000600004-1.txt

S0365-66912010000700004-1.txt

S0365-66912011000300005-1.txt

S0365-66912011000400005-1.txt

S0365-66912011001000004-1.txt

S0365-66912011001100006-1.txt

S0365-66912012000300003-1.txt

S0365-66912012000700004-1.txt

S0365-66912012000800004-1.txt

S0376-78922009000100005-1.txt

S0376-78922011000200004-1.txt

S0376-78922015000200010-1.txt

S0378-48352004000300006-1.txt

S0378-48352005000100005-1.txt

S0378-48352005000700007-2.txt

S0378-48352005000900005-1.txt

S0378-48352006000300006-1.txt

S0378-48352006000400005-1.txt

S0378-48352006000600006-1.txt

S1130-01082005001200011-1.txt

S1130-01082006000700014-1.txt

S1130-01082007000100013-1.txt

S1130-01082007000300006-7.txt

S1130-01082007000600012-1.txt

S1130-01082007000700016-1.txt

S1130-01082007001200014-1.txt

S1130-01082008000100010-1.txt

S1130-01082008000100017-1.txt

S1130-01082008000300010-1.txt

S1130-01082008000500014-1.txt

S1130-01082008000500017-2.txt

S1130-01082008000800011-1.txt

S1130-01082008000800019-1.txt

S1130-01082008000900011-1.txt

S1130-01082008000900015-1.txt

S1130-01082008001000010-1.txt

S1130-01082009000200011-1.txt

S1130-01082009000400012-1.txt

S1130-01082009000700018-1.txt

S1130-01082009000900012-1.txt

S1130-01082009000900015-1.txt

S1130-01082010000100021-1.txt

S1130-01082010000700014-1.txt

S1130-01082010001100014-1.txt

S1130-05582004000400006-1.txt

S1130-05582004000500005-1.txt

S1130-05582007000500006-1.txt

S1130-05582007000600006-1.txt

S1130-05582008000300004-1.txt

S1130-05582008000300005-1.txt

S1130-05582008000500007-1.txt

S1130-05582012000300006-1.txt

S1130-05582013000200006-1.txt

S1130-05582014000400006-2.txt

S1130-05582015000100004-1.txt

S1130-05582015000300007-1.txt

S1130-05582017000100031-1.txt

S1130-05582017000200122-1.txt

S1130-14732005000200003-1.txt

S1130-14732005000400006-1.txt

S1130-14732005000400007-1.txt

S1130-14732005000500006-1.txt

S1130-14732005000600007-1.txt

S1130-14732006000200011-1.txt

S1130-14732006000300007-1.txt

S1130-14732008000500005-2.txt

S1130-14732009000300006-1.txt

S1130-63432013000500014-1.txt

S1130-63432013000500017-1.txt

S1130-63432014000100011-1.txt

S1130-63432014000400016-1.txt

S1130-63432015000400006-1.txt

S1130-63432015000500008-1.txt

S1130-63432016000600013-1.txt

S1134-80462005000100007-1.txt

S1134-80462005000300007-1.txt

S1134-80462009000100005-1.txt

S1134-80462009000200004-1.txt

S1134-80462015000300003-1.txt

S1135-76062009000200005-1.txt

S1135-76062012000300014-1.txt

S1135-76062013000200010-1.txt

S1135-76062014000100006-1.txt

S1135-76062014000300006-1.txt

S1137-66272006000500011-1.txt

S1137-66272006000500012-1.txt

S1137-66272008000500007-1.txt

S1137-66272009000300013-1.txt

S1137-66272009000500016-1.txt

S1137-66272010000300010-1.txt

S1137-66272011000100013-1.txt

S1137-66272013000300020-1.txt

S1137-66272014000100021-1.txt

S1137-66272014000300015-1.txt

S1137-66272014000300016-1.txt

S1137-66272014000300017-1.txt

S1139-76322009000500006-1.txt

S1139-76322009000700016-1.txt

S1139-76322010000300008-1.txt

S1139-76322011000300007-2.txt

S1139-76322011000300008-1.txt

S1139-76322013000500010-1.txt

S1139-76322014000400007-1.txt

S1139-76322015000300013-1.txt

S1139-76322015000300018-1.txt

S1139-76322015000300018-2.txt

S1139-76322015000400016-1.txt

S1139-76322016000300012-1.txt

S1139-76322016000300013-1.txt

S1139-76322017000100010-1.txt

S1139-76322017000100011-1.txt

S1698-44472004000100009-1.txt

S1698-44472004000400011-1.txt

S1698-44472005000100002-1.txt

S1698-69462006000100019-1.txt

S1698-69462006000200016-1.txt

S1698-69462006000400005-1.txt

S1699-695X2014000200012-1.txt

S1699-695X2014000300011-1.txt

S1699-695X2015000100011-1.txt

S1699-695X2017000100011-1.txt

S1887-85712013000200013-1.txt

S1888-75462014000200009-1.txt

S1889-836X2015000200005-2.txt