The Intersection Inference Problem is NP-Complete

Mathman

November 25, 2009

First let me start off by defining the intersection inference problem.

Given a set U, some finite number ($n \in \mathbb{N}$) subsets A_i, and n constants c_i, is there a subset $X \subseteq U$ such that $|X \cap A_i| = c_i$ for i ranging all n values?

First we prove that intersection inference is in NP. Suppose we have a solution $X \subseteq U$. In order to verify that X is indeed a solution, we perform the m intersections of X with A and confirm that $|X \cap A_i| = c_i$. Set intersection can be performed (naïvely) in $\mathcal{O}(|X| \cdot |A_i|)$ time, so this process is indeed polynomial and clearly yields the correct answer. Thus intersection inference is in NP.

To prove intersection inference is NP-complete, we will show a reduction from 3-dimensional mapping to the intersection inference problem. With an arbitrary instance of the 3-dimensional mapping problem defined as sets X, Y, Z, T such that $X \cap Y = X \cap Z = Y \cap Z = \varnothing$, $|X| = |Y| = |Z| = n \in \mathbb{N}$ and $T \subseteq X \times Y \times Z$. To reduce this to the intersection inference problem, let $U = T$, and define $\{A_1, A_2, ..., A_n\}$ to be the subsets of triples that contain $x_1, x_2, ..., x_n \in X$ respectively. Define in a similar manner from $n+1$ to $2n$ for $y_i \in Y$ and from $2n+1$ to $3n$ for $z_i \in Z$. Finally define $A_0 = T$. The constraints are $c_0 = n$ and $c_i = 1$ for $i \in \{1, 2, ..., n\}$. We can see that to create these subsets takes $O(n^3)$ time in the rather straightforward manner of iterating over T's elements to divvy up the elements into arrays representing A_i.

Through this construction we can see that exactly n triples must match, and only one triple can contain each specific element in $X \cup Y \cup Z$. Thus if there exists a solution to this intersection inference problem, the solution is exactly the n triples that satisfy the 3-dimensional matching problem. If there does not exist a solution, then at least one constraint was not satisfied, meaning at least one element in $X \cup Y \cup Z$ was not covered, or all elements were covered with some redundancies. Thus the reduction is correct and intersection inference is NP-complete because a known NP-complete problem was poly-time reducible to the intersection inference problem which is in NP.