A Consensus Protocol in a Prison

Jayadev Misra

2/2/2004

The following problem was posed at a recent conference by Leslie Lamport, who had heard that it was posed in a popular radio program, *Car Talk*, in the National Public Radio in U.S. I solved a special case of this problem (assuming a certain initial state). Anup Rao, a student in my graduate class, solved the general case in an hour.

Problem A set of prisoners —assume there are at least 2— are asked to play the following game by the warden. There is a room in the prison which has two switches; initially, the switches are in arbitrary positions. The warden will bring one prisoner at a time to the room, and the prisoner must flip one of the switches. The prisoners do not know the order in which they will be taken to the room, but they know that every prisoner will visit the room over and over until the end of the game. The game ends when some prisoner announces, "every prisoner has been in this room at least once". If the announcement is correct, all prisoners go free; if incorrect, they all are executed. The game continues until the announcement is made.

The prisoners are allowed to confer and decide on a protocol prior to the start of the game. Once the game starts, they are not allowed to communicate, nor can they find out who is being taken to the room. The problem is to devise a protocol for the prisoners.

Solution; special case In the original formulation, there are two switches, A and B. Simplify the problem by assuming that there is a single switch S which a prisoner may or may not flip. This is a special case of the original formulation: whenever S is flipped, flip A, and whenever S is not flipped flip B. Henceforth, *flip* denotes flipping switch S. Let variable s denote the state of the switch, 0 for off and 1 for on; so, *flip* changes the value of s.

The prisoners choose a *leader* who will make the announcement; all other prosoners are *followers*. The number of followers, N, is at least 1. Each follower will vote by flipping the switch appropriately, and the leader will count the number of votes. The i^{th} follower has a variable v_i which is 1 if he has voted and 0 otherwise. Clearly, $v_i = 1$ implies that the i^{th} follower has visited the room. The leader has a variable c for maintaining the count. Initially, no follower has voted and c is 0. Assume, for the moment, that the switch is initially off; so, $s = 0$. The protocol is described by the steps taken by the leader and the followers.

---

initially: for all followers, $v_i := 0$; $c := 0$; $s := 0$

Follower **if** $v_i = 0$ {not voted} and $s = 0$ {switch is off} **then**
flip $\{s = 1\}$; $v_i := v_i + 1$
fi

Leader **if** $s = 1$ {switch is on} **then**
flip $\{s = 0\}$; $c := c + 1$; **if** $c = N$ **then** Announce **fi**
fi

The proof obligations are:

- (Safety) Whenever the leader announces, all v_is are 1.
- (Progress) Eventually, the leader announces.

Proof of Safety: Let v denote the sum of all the v_is. Then, $0 \leq v \leq N$. It is easy to see that $c + s = v$ is an invariant: it holds initially, and after each step by a follower or the leader. Since $c = N \ \wedge \ s = 0$ is a precondition of Announce, we conclude from $c+s = v$ that $v = N$ holds prior to Announce. Hence, all v_is are 1.

Proof of Progress: We show that c increases eventually if it is below N. Therefore, eventually $c = N$. Since c never decreases, $c = N$ will remain *true* and the leader will Announce. If $c < N \ \wedge \ s = 1$ then on its next visit to the room, the leader will increase c. If $c < N \ \wedge \ s = 0$ some follower has not voted yet; so, s will become 1 either by the voting of this or some other follower. Then, $c < N \ \wedge \ s = 1$ holds, and, from the previous argument, c will be increased.

Solution; general case We have assumed that initially $s = 0$. Now, we remove this assumption. Then, we have the invariant $c + s \leq v + 1$.

It is no longer true that $c = N \ \Rightarrow \ v = N$; v could be $N - 1$. However, c can be off by no more than 1 from v. This discrepancy can be handled *if every follower votes twice*. Now, the precondition of voting is changed to $v_i < 2$, and of Announce to $c = 2N$.

Proof of Safety: A precondition of Announce is $c = 2N \ \wedge \ s = 0$. From the invariant $c+s \leq v+1$, we get $2N \leq v+1$ as a precondition of Announce. We claim that then every prisoner has visited the room at least once, because if $N-1$ prisoners have visited the room and each has voted twice $v \leq 2N-2$, or $v+1 < 2N$.

Proof of Progress: Similar to the previous case.