

Function: Forward-phase(policy π)

1. Set *current_state* = START_STATE
2. **while**(*current_state* \neq FINAL_STATE)
3. Execute robot action *a* according to current policy π
4. Observe human action
5. Set *next_state* to the state resulting from *current_state*, robot and human action
6. Record *current_state*, *a*, *next_state*
7. *current_state* = *next_state*
8. **end while**