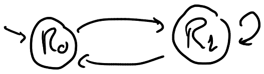
- A robot can move between 2 rooms R_0 and R_1 as follows:



i.e., (a) it starts in room R_0
(b) when in room R_0 it can only move to room R_1
(c) when in room R_1 it can move to R_1 or R_0

- To model this system, let us consider the following atomic propositions: $p_0, p_1, p_2, \ldots$
  such that p_i means that the robot is in room R_0 at step i (i.e., after i moves)
  and $\neg p_i$ means that the robot is in room R_1 at step i
- we can model the initial state of the robot as follows: p_0, i.e., the robot is initially in room R_0 ((a))
- we can model a transition as follows: $(p_i \rightarrow \neg p_{i+1}) \wedge (\neg p_i \rightarrow (\neg p_{i+1} \vee p_i))$, i.e., (b) $\wedge$ (c)
- we can capture two moves as follows: $\underbrace{p_0}_{\text{initial position}} \wedge \underbrace{(p_0 \rightarrow \neg p_1) \wedge (\neg p_0 \rightarrow (\neg p_1 \vee p_2))}_{\text{1st move}} \wedge \underbrace{(p_1 \rightarrow \neg p_2) \wedge (\neg p_1 \rightarrow (p_2 \vee p_2))}_{\text{2nd move}}$

- Could we prove that when in room R_1, the robot will be in room R_1 next? ← call it P
  No: when in R_1, the robot can move to R_0
  How do we formally prove that this property fails? we prove its negation, i.e. $\neg(\neg p_i \rightarrow \neg p_{i+1})$
- let's show that this fails after 2 steps, i.e.: $p_0 \wedge (p_0 \rightarrow \neg p_1) \wedge (\neg p_0 \rightarrow (\neg p_1 \vee p_1)) \wedge (p_1 \rightarrow \neg p_2) \wedge (p_1 \rightarrow (\neg p_2 \vee p_2))$
  $\wedge \neg(\neg p_1 \rightarrow \neg p_2)$

- We'll show this using a SAT solver: (1) convert the formula to a CNF using logical equivalences
  (2) use DPLL to check whether it is satisfiable

(1) $p_0 \wedge (p_0 \rightarrow \neg p_1) \wedge (\neg p_0 \rightarrow (\neg p_1 \vee p_1)) \wedge (p_1 \rightarrow \neg p_2) \wedge (\neg p_1 \rightarrow (\neg p_2 \vee p_2)) \wedge \neg(\neg p_1 \rightarrow \neg p_2)$

$\Leftrightarrow p_0 \wedge (\neg p_0 \vee \neg p_1) \wedge (\neg\neg p_0 \vee \neg p_1 \vee p_1) \wedge (\neg p_1 \vee \neg p_2) \wedge (\neg\neg p_1 \vee \neg p_2 \vee p_2) \wedge \neg(\neg\neg p_1 \vee \neg p_2)$ – elim $\rightarrow$

$\Leftrightarrow p_0 \wedge (\neg p_0 \vee \neg p_1) \wedge (\neg\neg p_0 \vee \neg p_1 \vee p_1) \wedge (\neg p_1 \vee \neg p_2) \wedge (\neg\neg p_1 \vee \neg p_2 \vee p_2) \wedge \neg\neg\neg p_1 \wedge \neg\neg p_2$ – de-Morg.

$\Leftrightarrow p_0 \wedge (\neg p_0 \vee \neg p_1) \wedge (p_0 \vee \neg p_1 \vee p_1) \wedge (\neg p_1 \vee \neg p_2) \wedge (p_1 \vee \neg p_2 \vee p_2) \wedge \neg p_1 \wedge p_2$ – $\neg\neg$ elim

CNF ↗

(2) use DPLL: $\cancel{p_0} \wedge \cancel{(\neg \cancel{p_0} \vee \neg p_1)} \wedge \cancel{(p_0 \vee \neg p_1 \vee p_1)} \wedge \cancel{(\neg p_1 \vee \neg p_2)} \wedge \cancel{(\cancel{p_1} \vee \neg p_2 \vee p_2)} \wedge \cancel{\neg p_1} \wedge \cancel{p_2}$

$p_0 = T$ $p_1 = F$ $p_2 = T$: the formula is satisfiable

This gives us a run of our robot: $R_0 \rightarrow R_1 \rightarrow R_0$, which shows that the property P does not hold, i.e., we obtained a counterexample.