

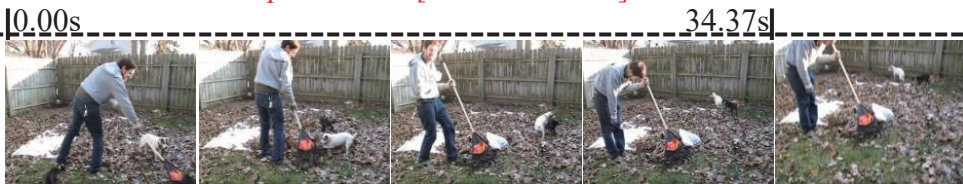


Question : What does the man do after he mashes the potatoes?

Ground Truth : He pours in the milk from the saucepan into the pot.

*Generated Answer : He pours milk.*

*Generated Visual Explanation : [140.23, 180.24]*

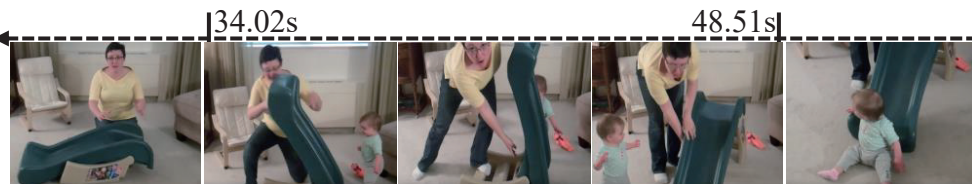


Question: What does the man do before bending down?

Ground Truth: He rakes the leaves.

*Generated Answer: : He sweeps the leaves.*

*Generated Visual Explanation : [0, 35.03]*

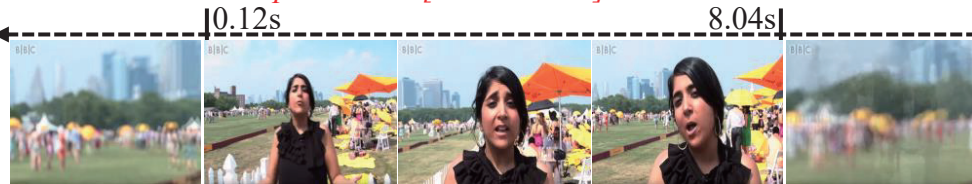


Question : What does the woman do while standing?

Ground Truth : She opens the holding bend slide in the living room.

*Generated Answer : She plays slide.*

*Generated Visual Explanation : [34.42, 47.98]*



Q: What does the woman wearing black clothes do when she first appears?

Ground Truth : She talks to the camera.

*Generated Answer: She talks to the camera.*

*Generated Visual Explanation: [0, 8.08]*