Ant

This environment is part of the Mujoco environments. Please read that page first for general information.

Action Space	Box(-1.0, 1.0, (8,), float32)
Observation Shape	(27,)
Observation High	[inf inf inf inf inf inf inf inf inf inf inf inf inf inf inf inf inf inf inf inf inf inf inf inf inf inf inf]
Observation Low	[-inf -inf -inf -inf -inf -inf -inf -inf -inf -inf -inf -inf -inf -inf -inf -inf -inf -inf -inf -inf -inf -inf -inf -inf -inf -inf -inf]
Import	`gym.make("Ant-v4")`

Description

This environment is based on the environment introduced by Schulman, Moritz, Levine, Jordan and Abbeel in "High-Dimensional Continuous Control Using Generalized Advantage Estimation". The ant is a 3D robot consisting of one torso (free rotational body) with four legs attached to it with each leg having two links. The goal is to coordinate the four legs to move in the forward (right) direction by applying torques on the eight hinges connecting the two links of each leg and the torso (nine parts and eight hinges).

Action Space

The action space is a `Box(-1, 1, (8,), float32)`. An action represents the torques applied at the hinge joints.

Num	Action	Control Min	Control Max	Name (in corresponding XML file)	Joint	Unit
0	Torque applied on the rotor between the torso and front left hip	-1	1	hip_1 (front_left_leg)	hinge	torque (N m)
1	Torque applied on the rotor between the front left two links	-1	1	angle_1 (front_left_leg)	hinge	torque (N m)
2	Torque applied on the rotor between the torso and front right hip	-1	1	hip_2 (front_right_leg)	hinge	torque (N m)
3	Torque applied on the rotor between the front right two links	-1	1	angle_2 (front_right_leg)	hinge	torque (N m)
4	Torque applied on the rotor between the torso and back left hip	-1	1	hip_3 (back_leg)	hinge	torque (N m)
5	Torque applied on the rotor between the back left two links	-1	1	angle_3 (back_leg)	hinge	torque (N m)
6	Torque applied on the rotor between the torso and back right hip	-1	1	hip_4 (right_back_leg)	hinge	torque (N m)
7	Torque applied on the rotor between the back right two links	-1	1	angle_4 (right_back_leg)	hinge	torque (N m)

Observation Space

Observations consist of positional values of different body parts of the ant, followed by the velocities of those individual parts (their derivatives) with all the positions ordered before all the velocities.

By default, observations do not include the x- and y-coordinates of the ant's torso. These may be included by passing `exclude_current_positions_from_observation=False` during construction. In that case, the observation space will have 113 dimensions where the first two dimensions represent the x- and y- coordinates of the ant's torso. Regardless of whether `exclude_current_positions_from_observation` was set to true or false, the x- and y-coordinates of the torso will be returned in `info` with keys `"x_position"` and `"y_position"`, respectively.

However, by default, an observation is a `ndarray` with shape `(111,)` where the elements correspond to the following:

Num	Observation	Min	Max	Name (in corresponding XML file)	Joint	Unit
0	z-coordinate of the torso (centre)	-Inf	Inf	torso	free	position (m)
1	x-orientation of the torso (centre)	-Inf	Inf	torso	free	angle (rad)
2	y-orientation of the torso (centre)	-Inf	Inf	torso	free	angle (rad)
3	z-orientation of the torso (centre)	-Inf	Inf	torso	free	angle (rad)
4	w-orientation of the torso (centre)	-Inf	Inf	torso	free	angle (rad)
5	angle between torso and first link on front left	-Inf	Inf	hip_1 (front_left_leg)	hinge	angle (rad)
6	angle between the two links on the front left	-Inf	Inf	ankle_1 (front_left_leg)	hinge	angle (rad)
7	angle between torso and first link on front right	-Inf	Inf	hip_2 (front_right_leg)	hinge	angle (rad)
8	angle between the two links on the front right	-Inf	Inf	ankle_2 (front_right_leg)	hinge	angle (rad)
9	angle between torso and first link on back left	-Inf	Inf	hip_3 (back_leg)	hinge	angle (rad)
10	angle between the two links on the back left	-Inf	Inf	ankle_3 (back_leg)	hinge	angle (rad)
11	angle between torso and first link on back right	-Inf	Inf	hip_4 (right_back_leg)	hinge	angle (rad)
12	angle between the two links on the back right	-Inf	Inf	ankle_4 (right_back_leg)	hinge	angle (rad)
13	x-coordinate velocity of the torso	-Inf	Inf	torso	free	velocity (m/s)
14	y-coordinate velocity of the torso	-Inf	Inf	torso	free	velocity (m/s)
15	z-coordinate velocity of the torso	-Inf	Inf	torso	free	velocity (m/s)
16	x-coordinate angular velocity of the torso	-Inf	Inf	torso	free	angular velocity (rad/s)
17	y-coordinate angular velocity of the torso	-Inf	Inf	torso	free	angular velocity (rad/s)
18	z-coordinate angular velocity of the torso	-Inf	Inf	torso	free	angular velocity (rad/s)
19	angular velocity of angle between torso and front left link	-Inf	Inf	hip_1 (front_left_leg)	hinge	angle (rad)
20	angular velocity of the angle between front left links	-Inf	Inf	ankle_1 (front_left_leg)	hinge	angle (rad)
21	angular velocity of angle between torso and front right link	-Inf	Inf	hip_2 (front_right_leg)	hinge	angle (rad)
22	angular velocity of the angle between front right links	-Inf	Inf	ankle_2 (front_right_leg)	hinge	angle (rad)
23	angular velocity of angle between torso and back left link	-Inf	Inf	hip_3 (back_leg)	hinge	angle (rad)
24	angular velocity of the angle between back left links	-Inf	Inf	ankle_3 (back_leg)	hinge	angle (rad)
25	angular velocity of angle between torso and back right link	-Inf	Inf	hip_4 (right_back_leg)	hinge	angle (rad)
26	angular velocity of the angle between back right links	-Inf	Inf	ankle_4 (right_back_leg)	hinge	angle (rad)

The remaining 14*6 = 84 elements of the observation are contact forces (external forces - force x, y, z and torque x, y, z) applied to the center of mass of each of the links. The 14 links are: the ground link, the torso link, and 3 links for each leg (1 + 1 + 12) with the 6 external forces.

The (x,y,z) coordinates are translational DOFs while the orientations are rotational DOFs expressed as quaternions. One can read more about free joints on the Mujoco Documentation.

Note: Ant-v4 environment no longer has the following contact forces issue. If using previous Humanoid versions from v4, there have been reported issues that using a Mujoco-Py version > 2.0 results in the contact forces always being 0. As such we recommend to use a Mujoco-Py version < 2.0 when using the Ant environment if you would like to report results with contact forces (if contact forces are not used in your experiments, you can use version > 2.0).

Rewards

The reward consists of three parts:

- *healthy_reward*: Every timestep that the ant is healthy (see definition in section "Episode Termination"), it gets a reward of fixed value `healthy_reward`
- *forward_reward*: A reward of moving forward which is measured as *(x-coordinate before action - x-coordinate after action)/dt*. *dt* is the time between actions and is dependent on the `frame_skip` parameter (default is 5), where the frametime is 0.01 - making the default *dt = 5 * 0.01 = 0.05*. This reward would be positive if the ant moves forward (in positive x direction).
- *ctrl_cost*: A negative reward for penalising the ant if it takes actions that are too large. It is measured as `ctrl_cost_weight` * sum(*action*^{2}) where `ctr_cost_weight` is a parameter set for the control and has a default value of 0.5.
- *contact_cost*: A negative reward for penalising the ant if the external contact force is too large. It is calculated `contact_cost_weight` * sum(*clip(external contact force to* `contact_force_range`*)*^{2}).

The total reward returned is ***reward*** *= healthy_reward + forward_reward - ctrl_cost - contact_cost* and `info` will also contain the individual reward terms.

Starting State

All observations start in state (0.0, 0.0, 0.75, 1.0, 0.0 ... 0.0) with a uniform noise in the range of [`-reset_noise_scale, reset_noise_scale`] added to the positional values and standard normal noise with mean 0 and standard deviation `reset_noise_scale` added to the velocity values for stochasticity. Note that the initial z coordinate is intentionally selected to be slightly high, thereby indicating a standing up ant. The initial orientation is designed to make it face forward as well.

Episode End

The ant is said to be unhealthy if any of the following happens:

1. Any of the state space values is no longer finite
2. The z-coordinate of the torso is **not** in the closed interval given by `healthy_z_range` (defaults to [0.2, 1.0])

If `terminate_when_unhealthy=True` is passed during construction (which is the default), the episode ends when any of the following happens:

1. Termination: The episode duration reaches a 1000 timesteps
2. Truncation: The ant is unhealthy

If `terminate_when_unhealthy=False` is passed, the episode is ended only when 1000 timesteps are exceeded.

Arguments

No additional arguments are currently supported in v2 and lower.

```
env = gym.make('Ant-v2')
```

v3 and v4 take gym.make kwargs such as xml_file, ctrl_cost_weight, reset_noise_scale etc.

```
env = gym.make('Ant-v4', ctrl_cost_weight=0.1, ...)
```

Parameter	Type	Default	Description
`xml_file`	**str**	`"ant.xml"`	Path to a MuJoCo model
`ctrl_cost_weight`	**float**	`0.5`	Weight for *ctrl_cost* term (see section on reward)
`contact_cost_weight`	**float**	`5e-4`	Weight for *contact_cost* term (see section on reward)
`healthy_reward`	**float**	`1`	Constant reward given if the ant is "healthy" after timestep
`terminate_when_unhealthy`	**bool**	`True`	If true, issue a done signal if the z-coordinate of the torso is no longer in the `healthy_z_range`
`healthy_z_range`	**tuple**	`(0.2, 1)`	The ant is considered healthy if the z-coordinate of the torso is in this range
`contact_force_range`	**tuple**	`(-1, 1)`	Contact forces are clipped to this range in the computation of *contact_cost*
`reset_noise_scale`	**float**	`0.1`	Scale of random perturbations of initial position and velocity (see section on Starting State)
`exclude_current_positions_from_observation`	**bool**	`True`	Whether or not to omit the x- and y-coordinates from observations. Excluding the position can serve as an inductive bias to induce position-agnostic behavior in policies

Version History

- v4: all mujoco environments now use the mujoco bindings in mujoco>=2.1.3
- v3: support for gym.make kwargs such as xml_file, ctrl_cost_weight, reset_noise_scale etc. rgb rendering comes from tracking camera (so agent does not run away from screen)
- v2: All continuous control environments now use mujoco_py >= 1.50
- v1: max_time_steps raised to 1000 for robot based tasks. Added reward_threshold to environments.
- v0: Initial versions release (1.0.0)