Text highlighted in Green is new (introduced by the framework). Uncolored text was in the participant's original requirements.

1. TITLE: Domain Knowledge Integration into Learning Process

SYSTEM_REQUIREMENT: The system must incorporate domain knowledge into state and action representations, reward function design, and exploration strategies to enhance learning efficiency and realism.
IMPLEMENTATION_DETAILS:
* State variables should include features inspired by animal behavior and handcrafted robot motions
* Partial rewards should be weighted sums including domain-informed terms
* Exploration policy should favor low-energy actions
* Efficient estimators should infer unobservable state components
* Unsafe actions should be penalized or removed from the action set.

RATIONALE: Incorporating domain knowledge improves the alignment of the RL model with real-world constraints and enables more efficient training by guiding exploration.

2. TITLE: Mitigation of Undesired Incentivized Behaviors

SYSTEM_REQUIREMENT: The system must penalize or eliminate behaviors that deviate from safe and efficient locomotion while maintaining the desired task objectives.
IMPLEMENTATION_DETAILS:
* Penalize known undesired behaviors with a large negative reward
* Eliminate unsafe actions from the action space where feasible
* Monitor for excessive joint angles or extreme speed to avoid structural degradation.

RATIONALE: A poorly designed reward function can reinforce unintended behaviors such as excessive speed or unsafe movements, which may degrade long-term performance or cause damage.

3. TITLE: Dimensionality Reduction for High-Dimensional State Spaces

SYSTEM_REQUIREMENT: The system must apply dimensionality reduction techniques to combine correlated state components into compact representations without significant loss of information.
IMPLEMENTATION_DETAILS:
* Combine root orientation and root angular velocity into a single representation
* Apply dimensionality reduction techniques to project state variables into a lower-dimensional space.

RATIONALE: Reducing the dimensionality of state representations improves computational efficiency and prevents redundancy while maintaining critical information.

4. TITLE: Performance Evaluation Metrics and Acceptable Ranges

SYSTEM_REQUIREMENT: The system must evaluate performance using metrics such as collision rate, falls per trial, task completion rate, recovery time, undesired behavior frequency, energy consumption, and completion time, with acceptable ranges determined based on environmental factors and theoretical optima.

IMPLEMENTATION_DETAILS:
* Track metrics including:
    ** collision rate
    ** falls per trial
    ** completion rate
    ** recovery time
    ** undesired behaviors
    ** energy consumption
    ** completion time
* Adjust acceptable ranges based on terrain type and theoretical optima.
RATIONALE: Well-defined performance metrics ensure that the learned policy meets safety and efficiency requirements across different environments.

5. TITLE: Validation Against Traditional Controllers

SYSTEM_REQUIREMENT: The system must validate RL-generated actions by comparing cumulative rewards, task completion rates, falls per trial, and other performance metrics against those of traditional controllers.
IMPLEMENTATION_DETAILS:
* Evaluate RL policy and traditional controller across shared performance metrics
* Compare cumulative rewards, task completion rates, and falls per trial
* Conclude comparability or superiority based on quantitative performance assessment.
RATIONALE: Comparing RL-generated actions with traditional control methods provides a benchmark for ensuring the reliability and effectiveness of learned policies.

6. TITLE: Environmental Variability in Training and Testing

SYSTEM_REQUIREMENT: The training and testing process must include randomized variations in environmental parameters such as wind force, obstacle placement, and other external factors.
IMPLEMENTATION_DETAILS:
* Randomize environmental parameters such as wind force and obstacle positions
* Apply randomization per episode during simulation
* Incorporate diverse environmental conditions in real-world tests
RATIONALE: To ensure the RL policy is robust to real-world environmental changes, the training and testing environments must incorporate variability in environmental dynamics.

7. TITLE: Prioritization of High-Risk Scenarios

SYSTEM_REQUIREMENT: A probability distribution function must be implemented to quantify and prioritize high-risk situations based on their negative consequences, ensuring more frequent exposure to critical scenarios during training.
IMPLEMENTATION_DETAILS:
* Human experts identify high-risk situations
* Define function quantifying impact of each scenario
* Construct probability distribution prioritizing high-risk scenarios

* Incorporate prioritized scenarios in replay strategies
RATIONALE: The RL system must be exposed to high-risk situations during training to improve resilience and ensure safe operation in critical scenarios.

8. TITLE: Handling Sensor Noise and Corrupted Data

SYSTEM_REQUIREMENT: Sensor data used in training must be augmented with artificially introduced noise, including perturbations based on expected sensor errors, as well as simulated missing or corrupted values.
IMPLEMENTATION_DETAILS:
* Introduce perturbations based on expected sensor noise variance
* Simulate missing/corrupted values using unrealistic values (e.g. negative numbers for positive-only parameters)
* Incorporate noisy data into training datasets
RATIONALE: To enhance robustness, the system must be capable of dealing with sensor inaccuracies and missing data during training.

9. TITLE: Inference of Missing or Unreliable State Information

SYSTEM_REQUIREMENT: The system must use available observations, past actions, and estimation techniques (e.g., Kalman filter) to infer missing or uncertain state data.
IMPLEMENTATION_DETAILS:
* Use historical observations and recent actions to infer missing states
* Apply Kalman filter or other estimation methods when applicable
* Use statistical imputation methods (mean, median, mode) when appropriate
RATIONALE: The system must be capable of estimating missing or unreliable state information to ensure continuous and reliable decision-making.

10. TITLE: Incorporating Action Execution Errors in Training

SYSTEM_REQUIREMENT: Action execution errors must be simulated in training by applying noise to executed actions, using a probabilistic model that represents expected actuator deviations.
IMPLEMENTATION_DETAILS:
* Simulate actuator noise by applying probabilistic deviation to selected actions
* Use a noise distribution to determine error magnitude
* Incorporate execution errors in both simulation and real-world tests
RATIONALE: To improve robustness, the system must account for real-world execution errors such as actuator noise and power limitations.

11. TITLE: Adaptation to Execution Deviations

SYSTEM_REQUIREMENT: The system must dynamically adjust action selection based on execution outcomes, prioritizing actions with lower variance and compensating for deviations using approximate error models when available.
IMPLEMENTATION_DETAILS:
* Adjust future actions based on observed deviations
* Prioritize actions with lower execution variance

* Use approximate error models for deviation compensation when available
RATIONALE: The RL system must prevent small execution errors from compounding into instability by adapting to deviations in action execution.

12. TITLE: Improving Training Efficiency with Limited Data
SYSTEM_REQUIREMENT: The training process must incorporate efficient exploration strategies, data augmentation, memory replay, pretrained policies, and regularization techniques to improve learning efficiency and generalization.
IMPLEMENTATION_DETAILS:
* Use efficient exploration strategies
* Implement data augmentation
* Apply memory replay techniques
* Leverage pretrained policies
* Use regularization to mitigate overfitting
RATIONALE: To enhance generalization and learning efficiency in data-limited scenarios, the system must employ techniques that maximize learning from available data.

13. TITLE: Hyperparameter Optimization for Stability and Generalization
SYSTEM_REQUIREMENT: A grid search must be performed for critical hyperparameters, selecting values that optimize generalization and performance across diverse environments.
IMPLEMENTATION_DETAILS:
* Perform grid search on crucial hyperparameters
* Select values that improve generalization and performance
RATIONALE: To balance stability, generalization, and performance, hyperparameters must be carefully selected through systematic optimization.

14. TITLE: Evaluation Metrics for Robustness
SYSTEM_REQUIREMENT: Robustness evaluation must include success/failure rates across environments and performance degradation under perturbations, with thresholds set relative to human-controlled baselines.
IMPLEMENTATION_DETAILS:
* Measure success/failure rate across diverse environments
* Assess performance degradation under perturbations (e.g.
* actuator failures)
* Establish evaluation thresholds based on human-controlled benchmarks
RATIONALE: To ensure robustness, the system must be evaluated using well-defined metrics that measure performance across diverse conditions and under perturbations.