**MOnarCH Storyboards: Joyful Warden Technical Script**

Table of Contents

[Introduction 3](#_Toc350161115)

[Definitions 4](#_Toc350161116)

[Required Behaviors 4](#_Toc350161117)

[Required Primitive Behaviors 6](#_Toc350161118)

[Required Perception Components 7](#_Toc350161119)

[Basic Components 7](#_Toc350161120)

[Application-Related Components 7](#_Toc350161121)

[Required Interfaces 9](#_Toc350161122)

[Scenarios and Scene Sequences 9](#_Toc350161123)

[Scenario 1 – Scene 1 9](#_Toc350161124)

[Scenario 1 - Scene 2 10](#_Toc350161125)

[Scenario 1 – Scene 3a 10](#_Toc350161126)

[Scenario 1 – Scene 3b 11](#_Toc350161127)

[Scenario 1 – Scene 3c 11](#_Toc350161128)

[Scenario 2 – Scene 1 12](#_Toc350161129)

[Scenario 2 – Scene 2 12](#_Toc350161130)

[Scenario 2 – Scene 3 12](#_Toc350161131)

[Scenario 3 – Scene 1 13](#_Toc350161132)

[Scenario 3 – Scene 2a 13](#_Toc350161133)

[Scenario 3 – Scene 2b 13](#_Toc350161134)

[Scenario 4 – Scene 1 14](#_Toc350161135)

# Introduction

This document details the “Joyful Warden” storyboard. The goals for this storyboard are as follows:

* To demonstrate that a group of robots can help the staff at IPOL managing children in a closed environment (such as IPOL),
* To provide reasonable evidence (an unambiguous demonstration may be very difficult) that social relationships between children and robots can be established,
* To demonstrate that robots can be endowed with social interaction skills allowing them to mix with humans in social environments, using state-of-the-art techniques and off-the-shelf technologies.

One to two PO robots and two to three SO robots are considered.

The role of the PO robots is not explicit in this document. It is assumed that they behave without human-robot interaction concerns as their only objective is to obtain useful data for perception. However, they must behave according to the established social norms.

People participating in the MOnarCH trials are assumed to wear a wristband (or similar) with a RFID tag and eventually with some other sensing device.

The document is organized by definitions (section 1), required resources, both physical and functional (sections 2 to 5), and the script sequences (section 6).

The script is organized as sequences of scenes. Four possible initial scenarios are identified. For each of these scenarios, a full set of sequences of scenes is detailed. Continuity of the scenes must be ensured at all times to avoid having the system in deadlock/livelock situations.

# Definitions

**Role**: A set of behaviors all related to some specific task.

**Behavior**: An action on the environment (people included). It can lead to a robot moving, interacting through speech, etc. Compositions of behaviors to generate more complex behaviors are included here.

In general a behavior will be a composition of primitive behaviors.

**Primitive behavior**: An atomic behavior. Primitive behaviors are not intended to be decomposed into simpler behaviors. Examples are basic navigation, grasping, gestures, and voice dialogue.

**Perception component**: A provider of some perception feature. This includes algorithms running over either realtime data or data stored in some database that can be accessed concurrently by any component in the MOnarCH system.

**Interface**: A physical device (and associated drivers if appropriate) used to interact with the environment. A most natural interface is the robot itself as it interacts with the environment through motion. Other interfaces include speakers, LED arrays, etc.

# Required Behaviors

**B01 - Cooperative coverage of the environment**

Description: Move around the environment, in cooperation with the other robots, such that a complete coverage is assured. Eventually, the coverage algorithms for PO and SO robots may differ.

**B02 - Human-aware wandering**

Description: Wander around the environment accounting at all times for the social dynamics.

**B03 - Point-to-point human-aware navigation**

Description: Go from point A to point B without too much concern for rigid social rules. Only a “reasonable” set of rules needs to be accounted for, namely those related to safety.

**B04 - Staff member greeting**

Description: Greet a staff member if it detects a wristband with a RFID tag id corresponding to a staff member.

**B05 - Show warm expressiveness**

Description: Show a warm expression, either by adopting some adequate motion, displaying some happy facial expression, or using some other anthropomorphic feature.

**B06 - Get child RFID id**

Description: Approach a child to get his/her RFID id. The approaching motion may not be subject to the same rules usually used for humans, e.g., should not approach from behind.

**B07 - Get Staff member RFID id**

Description: Approach a staff member to get his/her RFID id. The approaching motion may be subject to specific rules, e.g., always approach from front.

**B08 - Safe voice interaction reply to child**

Description: The robot speaks to a child in reply to some speech interaction from the child. The sentences are tailored to minimize the complexity of an interactive speech dialogue.

**B09 - Safe voice interaction with child**

Description: The robot speaks to a child without any prior speech from the child. The sentences are tailored to minimize the complexity of an interactive speech dialogue.

**B10 - Greet a child**

Description: This is just a specialized form of the “safe voice interaction with child” behavior.

**B11 - Follow a child**

Description: The robot follows a child from a safe distance. Children tend to have more erratic movements than adults and hence a safety distance bigger than for adults may be used. This behavior is a specialized (more complex) form of the more general behavior “follow a human”.

**B12 - Follow a staff member**

Description: The robot follows a staff member from a safe distance. This behavior is also a specialized (more complex) form of the more general behavior “follow a human”.

**B13 - Outrun a child**

Description: The robot moves ahead of a child to a specific point or area in the environment. This behavior may imply some timing, and not only distance, constraints.

**B14 – Cooperative outrun of a child**

Description: Similar to B13 but instead of having a robot moving too fast to outrun a child it uses a cooperative strategy. Assuming that multiple robots are spread around the environment, the child is outrun by the most adequate robot, for instance the one that can move safer.

**B15 - Cooperative child following**

Description: The cooperative version of the “Follow child” behavior. One robot may act as leader of the team of robots. This leader robot may follow the child. The other(s) robot(s) will cooperate/coordinate with the leader to reach specific goals. Some of the follower robots may use the “outrun a child” behavior.

**B16 - Child-Robot Interaction navigation**

Description: The robot moves in the neighborhood of a single child using proxemics rules.

**B17 - Children-Robot Interaction navigation**

Description: The robot moves in the neighborhood of a group of children using proxemics rules.

**B18 - Staff-Robot Interaction navigation**

Description: The robot moves in the neighborhood of a member of the staff trying to get her/his attention.

**B19 – Move to block access to forbidden area**

Description: The robot moves to block the access of a child to a forbidden area.

**B20 – Cooperative movement to block access to a forbidden area**

Description: Similar to B19 but using multiple robots moving in cooperation to block the access to a forbidden area.

# Required Primitive Behaviors

**BP01 - Point-to-point asocial navigation**

Description: Go from point A to point B avoiding obstacles.

**BP02 - Issue spoken message**

Description: Issue a speech portion. The message may be configured to be spoken in multiple tones. For example, there must be similar primitive behaviors using warm and cold voice tones.

**BP03 - Follow a human**

Description: Follows a child, staff member, or a participant in the trials from some predefined distance. This distance is fixed whenever possible and may be modified if the environment conditions require it. Eventually, changes in this safety distance must have some correspondence with changes in the robot velocity.

# Required Perception Components

### Basic Components

**PC01 - Robot localization**

Description: Returns the position and orientation of each robot in the system in world coordinates and the associated uncertainty. It uses sensor fusion whenever possible, i.e., when the data from each sensor is both reliable and not too old.

**PC02 - Robot localization from fixed cameras**

Description: Returns the localization of a robot using the images from the camera network.

**PC03 - Robot localization from onboard LRF**

Description: Returns the localization of a robot from laser range finder measures, eventually using a SLAM strategy.

**PC04 - Robot localization from bar code**

Description: Returns the localization of a robot from the detection of bar codes placed on the walls.

**PC05 - Robot localization from RFID tags**

Description: Returns the localization of a robot from the detection of RFID tags placed on the walls.

**PC06 - Shell touching**

Description: Returns information on the touching of the outer shell of a robot by a child.

**PC07 - Speech detection and recognition**

Description: Recognition of, at least, a set of sentences useful for MOnarCH. Eventually an off-the-shelf package will be used for this purpose. Whatever the package, in case it can not perform the detection it must return state information that can be used by the system for decision making purposes, e.g., the sentence is not loud enough for recognition purposes, sentence only partially recognized, etc.

**PC08 – Unmapped environment features detection**

Description: Detects changes between the environment known to the system (eventually stored in a knowledge base) and the current environment.

### Application-Related Components

**P01 - Detection of children groups**

Description: Estimate the position (in world coordinates) of a group of children. Yet to be decided the best metric. Eventually a mass-center like metric may be enough.

**P02 - Detection of single children (not in groups)**

Description: Estimate the position (in world coordinates) of a single child in the environment. Yet to be decided the best metric. It may depend on the specific sensors used.

**P03 - Recognition of children group activities**

Description: Examples of interest for the storyboard are playing, wandering, running, and static. The corresponding perception component(s) must estimate some dynamics parameters, e.g., an estimate for the group motion velocity.

**P04 - Recognition of single children activities**

Description: Idem the previous component but for a single child.

**P05 - Toy tracking**

Description: Tracking of the position of a toy tagged with RFID.

**P06 - Get RFID id for human (child / staff member / participant)**

Description: Get the RFID tag info from the wristband from child or staff (or any other participant). Eventually it may be necessary to separate this component in 2 different (more specialized) components as follows

* P06a – Get child RFID id

Children are normally smaller than adults and this type of physical feature may constraint the detection of the tags.

* P06b – Get staff member RFID id

Adults may wear wristbands with the tags but also may have a tag directly placed on clothes.

**P07 - Track a child**

Description: Tracks the position of a single child using the sensor network.

**P08 – Detect abandoned toy**

Description: Use imaging and/or RFID detection to find abandoned toys. This task may be carried out by the PO robots.

**P09 - Child social behavior recognition**

Description: Recognition of social activities of children. Examples are “child playing”, “child running”, “child talking to adult”, “child talking to other child”, “child agitated”, “child sleeping”, etc

**P10 – Children social behavior recognition**

Description: Identical to P09 but for groups of children.

**P10 - Track children**

Description: Similar to P07 but for groups of children.

# Required Interfaces

**Motion**. This is an intrinsic interface, referred here just for the sake of completeness.

**Earing**. Microphones integrated within a speech recognition system.

**Voice**. Loudspeakers integrated within a dialogue management system.

**Display**. The physical installation of the display still to be defined by robot designers, e.g., as a face of the robot, as part of a torso, etc.

**Touch**. The shell of the robots may include touch sensors.

**Video projector**. This is part of the augmented interaction capabilities.

**LED arrays**. Thse may be used for facial expressions (mouth, eyebrows).

# Scenarios and Scene Sequences

Four scenarios are distinguished as starting scenes.

1. Scenario 1: Everything is quiet
2. Scenario 2: Single kids are wandering (not playing)
3. Scenario 3: Single kids are playing with bikes
4. Scenario 4: At least a group of kids is playing

Each of these scenarios is composed by multiple scenes. Figure 1 shows the estimated interconnections between the scenes.



Figure 1: Interconnections between the estimated scenarios/scenes

## Scenario 1 – Scene 1

Robots are either (i) quiet in some pre-defined areas or (ii) patrolling the corridor and playground room. Eventually, patrolling can be used by some of the PO robots to acquire information on the environment, i.e., re-check the consistency of environment map.

Children are behaving quietly not requiring any special attention from the robots.

If someone from the staff is seen for a first time in the day by any SO robot then the robot issues some greeting message.

This scenario can be followed by any other scene from any scenario.

Behaviors required: B03 (Point-to-point human-aware navigation). B04 (Staff member greeting). B05 (Show warm expressiveness). B01 (Cooperative coverage of the environment).

Perception required: P01 (Detection of children groups). P03 (Single children detection). PC08 (Unmapped environment features detection). P06a (Get child RFID id). P06b (Get staff member RFID id).

## Scenario 1 - Scene 2

A child moves out from a room and starts heading to the play room. If there the corridor is free then one SO robot may simply approach the child to first detect his/her RFID id. If it is someone participating in the MOnarCH trials it is greeted by the robot. Otherwise the robot moves away from the child without attempting any interaction. The other robots present will also not engage in any interaction.

If the child is participating in the trials, the robot then follows the child or stay in a close neighborhood. If the child attempts to interact with the robot through voice the robot issues a reply that does not foster any subsequent speech interactions (which may be difficult for the robot to handle), e.g., the robot can speak “I just like to watch you” in reply to a number of possible sentences from the child.

Behaviors required: B06 (Get child RFID id). B10 (Child greeting). B08 (Safe voice interaction reply to child). B11 (Follow a child).

Perception required: P04 (Recognition of single children activities). P02 (Detection of single children). PC07 (Speech detection and recognition).

There are multiple possible scenes that can follow scene 2. Let us define only three, which seem closer to the idea of a joyful warden, (i) the child ignores the robot, (ii) the child moves towards the “Garage” area in the playroom, and (iii) the child moves towards any other area.

## Scenario 1 – Scene 3a

This scene applies when the child ignores the robot.

The robot does not force any interaction with the child. It moves back, eventually following the child from a distance. Eventually, after a while the robot moves away from the child and returns to scenario 1, scene 1.

Behaviors required: B11 (Follow a child).

Perception required: P07 (Track a child).

## Scenario 1 – Scene 3b

This scene applies when the child moves towards the “Garage” area.

Once the robot knows that the child is approaching the “Garage” area, it issues some encouraging sentence, e.g, “Good we are going to play with the bikes”. It still keeps a safe distance from the child and waits for him/her to pick up the bike. In case that does not happen scenes 3a or 3c apply.

If the child picks up a bike and heads to the corridor the robot will assume that the child intends to run the bike in the corridor. Since at least two SO robots will be used, one of the robots may outrun the child and move straight to the beginning of the corridor while the second robot follows the child. Note that in general outrunning a child may not be possible (because of limitations of the robots), or reasonable (because it may require velocities too high to be safe). A cooperative outrun behavior may be used instead.

If the child aborts the bike run scenes 3a or 3c apply.

When the child reaches the corridor and one robot is already there it may be greeted by the robot. The robot following the child will join the other robot forming a group of attendees of the bike run. One of the robots uses the augmented HRI interfaces e.g., video projection, to stimulate the child. Examples of possible contents for video projection are (i) display images of children running bikes on a wall, (ii) display a “finish line” on the ground, (iii) display a start-run timer, etc.

While the child is running the bike the robots may be cheering. If the child wants to use too much space in the corridor one of the robots moves to block access to forbidden areas. At the same time the child is warned that in can not go to the forbidden areas.

Once the child stops running the bike, if he/she leaves the bike unattended one of the robots will issue a message requesting the child to park the bike in the proper place. If the child ignores the warnings one robot will follow him/her until the bike gets parked in the right place. The robot also issues a message to the child asking him/her to help accomplishing a task that the robot can not do by itself.

If the child ignores the robot but a parent is going to park the bike in the “Garage” one robot says “thanks” to the parent.

To simplify detection, bikes (and even other toys) may be RFID tagged.

Behaviors required: B14 (Cooperative child following). B08 (Safe voice interaction with child). B13 (Outrun a child). B14 (Cooperative outrun of a child). B19 (Move to block access to forbidden area). B20 (Cooperative movement to block access to forbidden area).

Perception required: P04 (Recognition of child activities). P05 (Toy tracking). P07 (Child tracking). P08 (Detection of abandoned toy).

## Scenario 1 – Scene 3c

This scene applies when the child moves towards any other area.

One robot follows the child. If he/she moves towards some forbidden area the robot will issue some warning. If he/she just approaches their parents or someone from staff the system will assume that the child is taken care by the parents or staff and will back off. The system is then in scenario 1, scene 1.

Otherwise the robot closer to the child will continue following the child (to convey the idea that the robots are thinking about what he/she is doing). The system is then in scenario 1, scene 2.

Behaviors required: B11 (Follow child). B09 (Safe voice interaction with child).

Perception required: P07 (Child tracking).

## Scenario 2 – Scene 1

Similar to Scenario 1 – Scene 1. Robots are either (i) quiet in some pre-defined areas or (ii) patrolling the corridor and playground room. Eventually, patrolling can be used by some of the PO robots to acquire information on the environment, i.e., re-check the consistency of environment map.

If a child wanders in a small area then robots do not disturb the child and the system is still in this scenario and scene.

If a child moves between larger areas, e.g., from the playroom to the corridor, then one robot will go searching for the child (scene 2).

Behaviors required: B02 (Human-aware wandering).

Perception required: P02 (Detection of single children).

## Scenario 2 – Scene 2

The robot that is closest to the area the child is heading to moves to approach the child. It keeps a safe distance without forcing interaction. Eventually some the robot can greet the child.

If the child ignores the robot then move to scenario 1 – scene 3a.

If the child initiates an interaction then move to scene 3.

Behaviors required: B03 (Point-to-point human-aware navigation). B11 (Child greeting).

Perception required: P02 (Detection of single children). P07 (Child tracking).

## Scenario 2 – Scene 3

This scene applies when a child starts an interaction with a robot.

The child may interact the robot using voice and touch. Voice interaction is managed by a dialogue management system. The dialogue must be consistent with the age of the child. The robot may tend to be a little inquisitive in order to minimize the possibility of the child asking questions to which the dialogue management system can not handle.

During the dialogue the robot will be facing the child most of the time.

If the child breaks the interaction the robot moves away returning to scene 1

After some time the robot may break the interaction arguing, for example, that it is tired and needs some rest returning to scene 1.

Behaviors required: B16 (Child-Robot Interaction navigation). B09 (Safe voice interaction with child).

Perception required: P02 (Detection of single children). P07 (Child tracking). PC07 (Speech detection and recognition). PC06 (Shell touching).

## Scenario 3 – Scene 1

Similar to Scenario 1 – Scene 1. Robots are either (i) quiet in some pre-defined areas or (ii) patrolling the corridor and playground room. Eventually, patrolling can be used by some of the PO robots to acquire information on the environment, i.e., re-check the consistency of environment map.

If all children detected are not moving to form groups then select the most agitated child playing and send one robot towards that child (scene 2a). Otherwise, if some children are moving to form groups select the group with the highest dynamics (scene 2b) and send one robot towards that group.

Behaviors required: B03 (Point to point human aware navigation).

Perception required: P02 (Detection of single children). P09 (Child social behavior recognition).

## Scenario 3 – Scene 2a

The robot approaches the agitated child carefully showing interest by the activity of the child. It queries the child on what he/she is doing using a sentence that does not foster subsequent interactions, e.g., let the child know that the robot likes to watch him/her playing.

If the child starts an interaction then scenario 2, scene 3, applies.

Behaviors required: B11 (Child greeting). B09 (Safe voice interaction with child). B16 (Child-Robot Interaction navigation).

Perception required: P02 (Detection of single children). P10 (Children social behavior recognition).

## Scenario 3 – Scene 2b

This scenario started with single children playing. However, given the limited amount of space it is likely that children tend to group rapidly and hence a group playing scene unfolds.

The robot approaches the group of children playing carefully showing interest by the activity of the children. It queries the group on what they are doing using a sentence that does not foster subsequent interactions to minimize complexity of dialogue management, e.g., let the group know that the robot likes to watch them playing.

If the children head to the corridor with the bikes then it is safe to assume that they intend to race the bikes in corridor. The same behaviors used in scenario 1, scene 3b are applicable here.

If the children abort the playing then scenario 1, scene 1, or scenario 2, scene 1 may apply.

Behaviors required: B09 (Safe voice interaction with children). B16 (Children-Robot Interaction navigation). B15 (Cooperative child following). B13 (Outrun a child). B14 (Cooperative outrun of a child).

Perception required: PC07 (Speech detection and recognition). P05 (Toy tracking). P10 (Children tracking). P03 (Recognition of children activities).

## Scenario 4 – Scene 1

One or two SO robots move towards the most entropic group of children playing.

As long as a group of children clearly stands out from the others the sequence of scenes is similar to that in scenario 3, for a single child.

If more than one group stands out there will be not enough robots to handle the situation. Each of the SO robots then chooses a group of children and follows then according the behaviors used to follow a single child at a distance in scenario 3.

Behaviors required: B02 (Point-to-point human aware navigation). B09 (Safe voice interaction with children). B16 (Children-Robot Interaction navigation). B15 (Cooperative child following). B13 (Outrun a child). B14 (Cooperative outrun of a child).

Perception required: P03 (Recognition of children group activities). P07 (Speech detection and recognition). P05 (Toy tracking). P10 (Children tracking).