Riccardo Caccavale

Tutor: Alberto Finzi

XXIX Cycle - III year presentation

Flexible Task Execution and Cognitive Control in Human-Robot Interaction

Background

Graduation MS: Computer Science (Computational Models) Federico II.

DIETI groups: PRISCA laboratory of advanced cognitive science, PRISMA laboratory of industrial robotics.

Collaborations: LAAS-CNRS (Toulouse), TUM (Munich)

Credit Summary

	Credits year 1								Credits year 2								Credits year 3									
	Estimated	bimonth 1	bimonth 2	bimonth 3	bimonth 4	bimonth 5	bimonth 6	Summary	Estimated	bimonth 1	bimonth 2	bimonth 3	bimonth 4	bimonth 5	bimonth 6	Summary	Estimated	bimonth 1	bimonth 2	bimonth 3	bimonth 4	bimonth 5	bimonth 6	Summary	Total	Check
Modules	18					3	14	**17**	15		3				3	**6**	7						6	**6**	29	30-70
Seminars	13				2,8	2,2		**5**	4	1				5,8		**6,8**	0							**0**	12	10-30
Research	34	10	10	8	8	7	5	**48**	45	9	7	10	10	5	7	**48**	40	8	8	8	8	8	3	**43**	139	80-140
	65	10	10	8	11	12	19	**70**	64	10	10	10	10	11	10	**61**	47	8	8	8	8	8	9	**49**	180	180

Period Abroad:

3/05 – 31/05 and 21/06 – 28/07 at TUM university of Munich.

Flexible Human-Robot Interaction

In social and industrial robotics **flexible** and **natural** interaction with humans is **needed** to perform structured collaborative tasks.

Robots should be able to **flexibly adapt** action execution to the human behaviors and the environmental changes.

Plan-based: relevant works in rely on the *three layer architectures* [Gat et al.] exploiting planning, re-planning [Ayan et al] and plan-adjustment [Estlin et al].

- Sometimes slow to react to unexpected events or human interventions.

Behavior-based: some approaches involve *reactive* and *behavior-based* systems [Arkin] to on-line adapt task/plan execution [Nicolescu et al, Proetzsch et al].

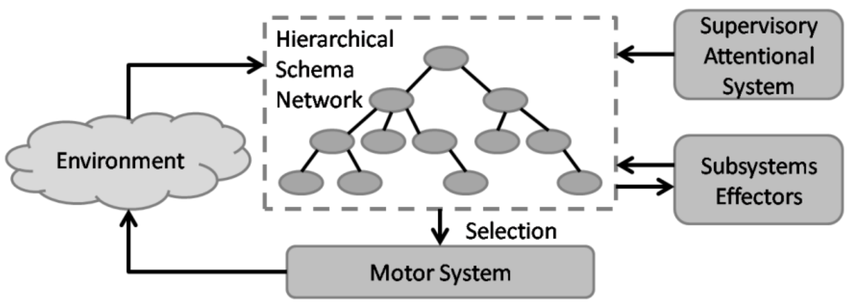
- Need to be orchestrated.

Cognitive-based: *cognitive architectures* [Langley et al] have also been proposed to develop flexible and human-friendly robotic frameworks [Anderson et al, Albus et al].

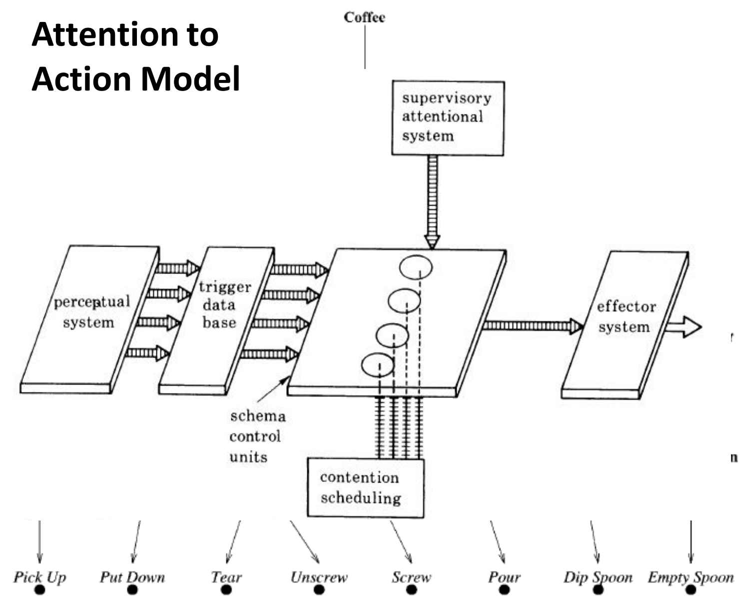
- Focused on psychology rather than robotics.

Cognitive Control and Attention

Cognitive control: *ability to configure itself for the performance of specific tasks through appropriate adjustements in perceptual selection, response biasing, and on-line maintenance of contextual information. [Botvinick et al. 2001]*

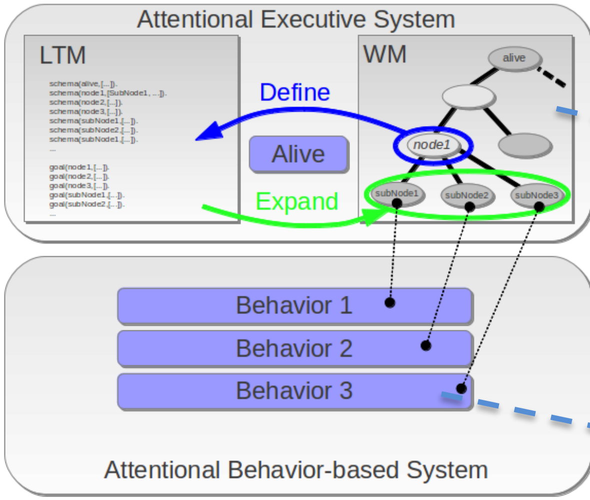


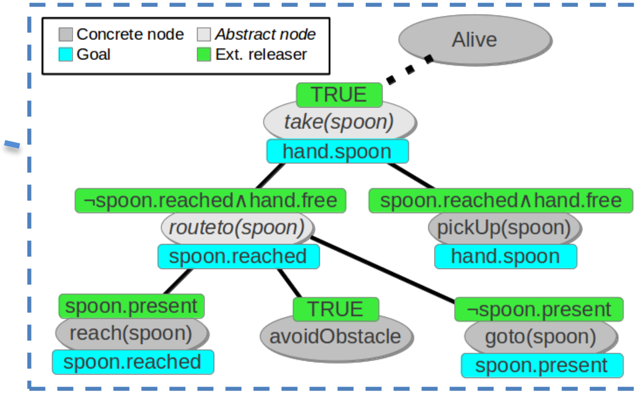
Attention to Action Model

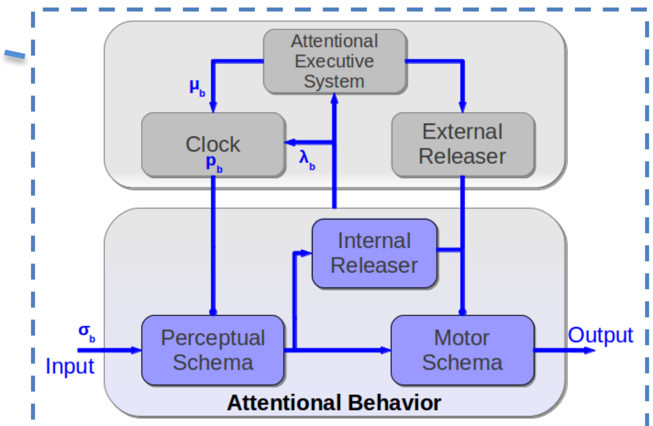


- **Contention Scheduling:** conflicting actions have to compete each other: only the most active actions are executed.
- **Supervisory Attentional System:** is employed to oversee the contention and regulate actions according to the context [Norman and Shallice 80].

Attentional Executive System







Working Memory: maintains the **executive state** and the structure of the tasks in the **attentional focus** of the system, including all the task the system is executing or willing to execute.

Long Term Memory: is a repository which contains the definition of all the **tasks available** to the system.

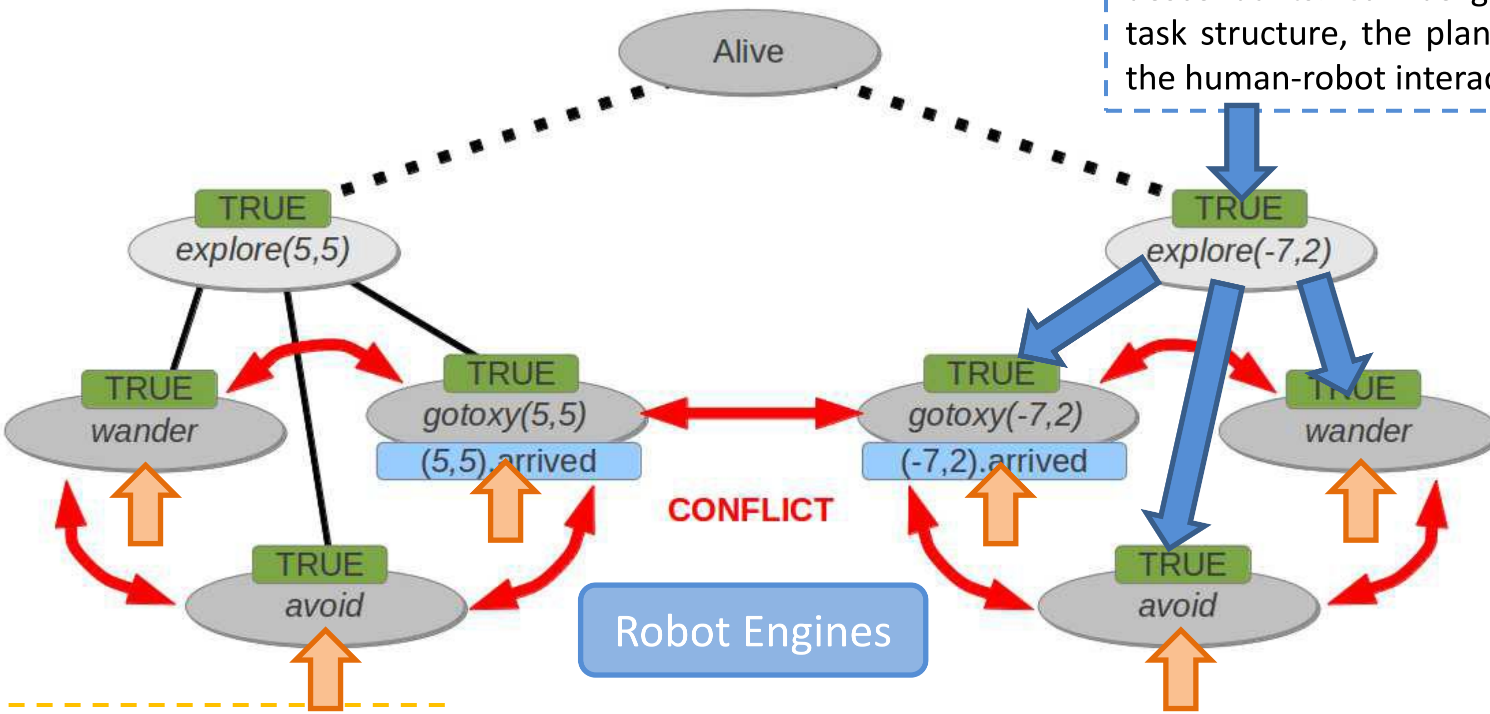
Attentional Behaviors: **elaborates** sensor data (behavior-specific stimuli) producing a pattern of **motor actions**.

Emphasis: **integrates** top-down (μ_b) and bottom-up (λ_b) stimulation to manage the **contention** between behaviors and **clocks**.

$$e_b = \mu_b \lambda_b$$

Contention and Regulations

Top-down: is inherited by all descendants. Can be given by the task structure, the plan or through the human-robot interaction.



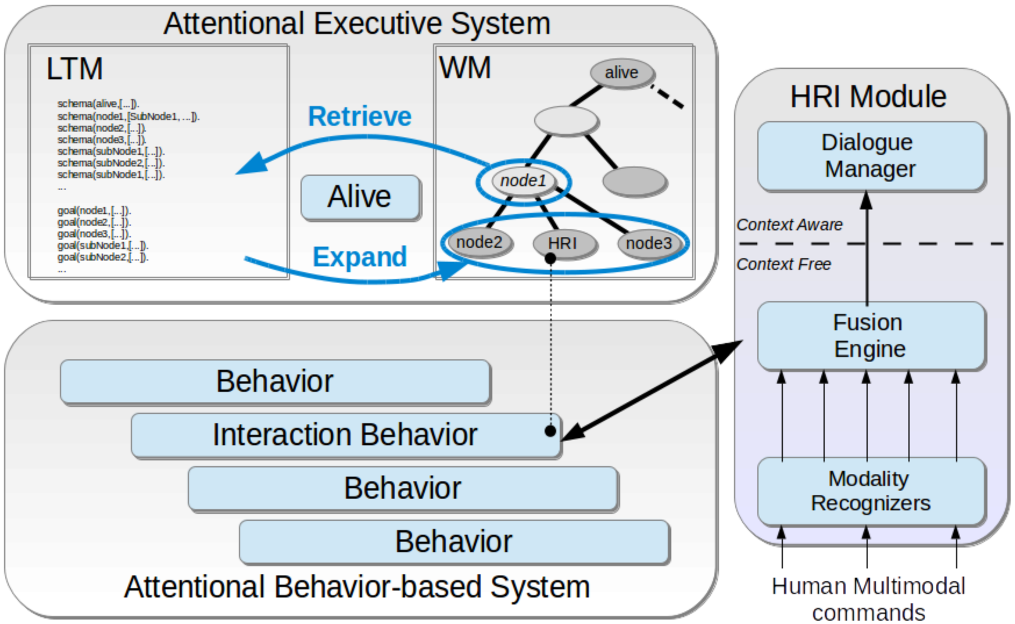
Bottom-up (stimuli-driven): given by salient elements of the environment (objects, landmarks, obstacles, etc.).

Contentions among behaviors competing for mutually exclusive state variables are solved following a *winner-takes-all* approach. The most emphasized behavior is selected with the exclusive access.

Summary

- In many robotic contexts a robotic systems should be able to dynamically execute activities, react to human behaviors and environmental changes.
- Prominent approaches rely on planning/re-planning processes during the task execution that are time consuming.
- In the proposed approach we exploit contention scheduling and supervisory attention during the task execution.
- The system includes attentional regulations mechanisms for the flexible orchestration of robotic sensorimotor processes.
- During this PhD, the proposed system has been applied in different robotic contexts including HRI, flexible plan execution, and learning…

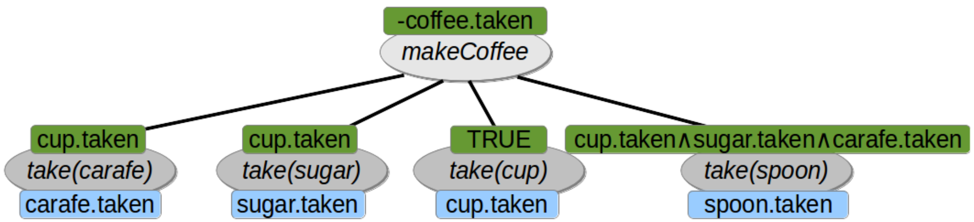
Attentional Regulations in HRI

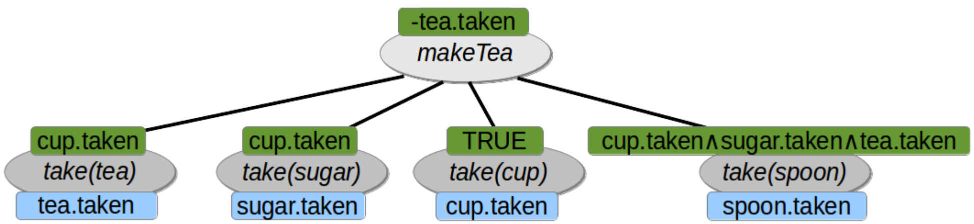


- Commands recognized by the HRI module are associated with tasks to be allocated in WM.
- Tasks can be instantiated with contextual subtasks and arguments, while their execution is regulated by top-down/bottom-up attentional mechanisms.

1. **Recognition layer** acquires the input from sensors and provide a classification for each interaction modality.
   1. Speech: Julius based recognition.
   2. Gestures: Latent Dynamic Conditional Random Fields (LDCRF).
2. **Fusion Layer** acquires the N-best gestures and speech from the classifiers to produce a list of N-best users intention.
3. **Dialogue Layer** integrates information about the dialogue history that is represented as a Partially Observable Markov Decision Process (POMDP) producing a user intention according to his policy.

Interaction Ambiguities: Coffee or Tea?



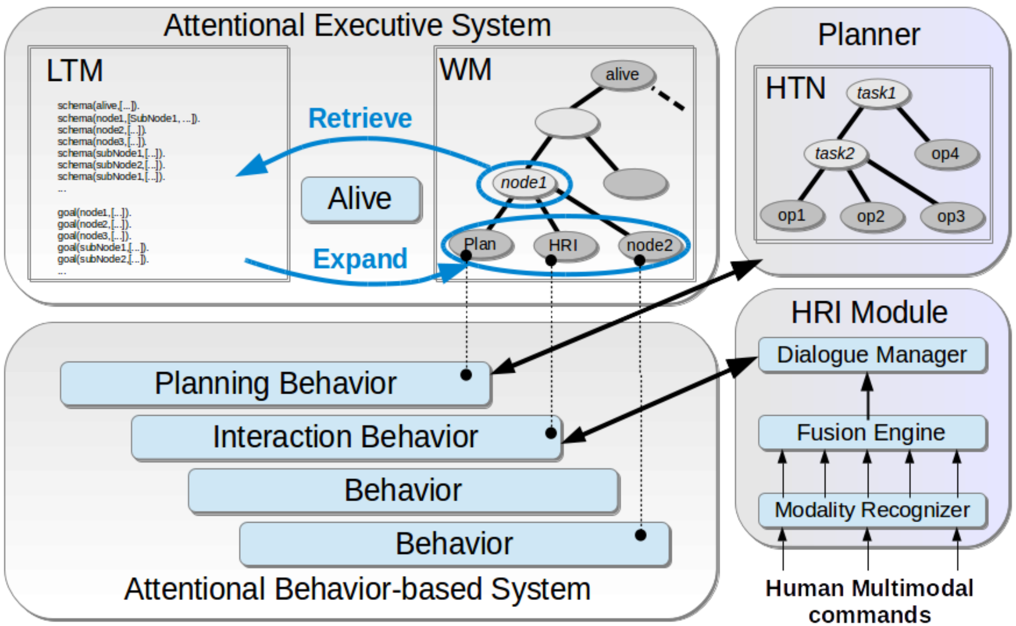


Interactive Commands: some user's acts are not explicit commands, therefore the system should interpret the human intention supporting the human activity with a proactive behavior.

Test Scenario: we consider the system at work in a kitchen scenario where two ambiguous tasks (make-coffee and make-tea) should be interpreted by the system.

	Success	Correction	Failures
Percent.	**56.6%**	26.7%	16.7%
Hum. Act.	1.48	2.25	3.6
Std.	0.67	0.42	0.52

Flexible Plan Execution



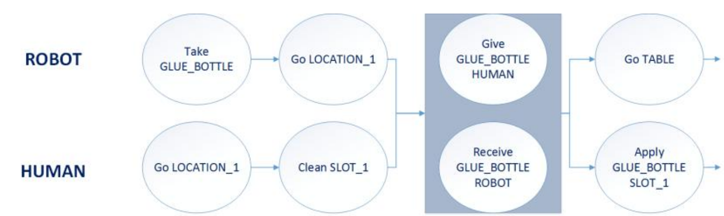
SAPHARI
SAFE AND AUTONOMOUS PHYSICAL HUMAN-AWARE ROBOT INTERACTION

SEVENTH FRAMEWORK PROGRAMME

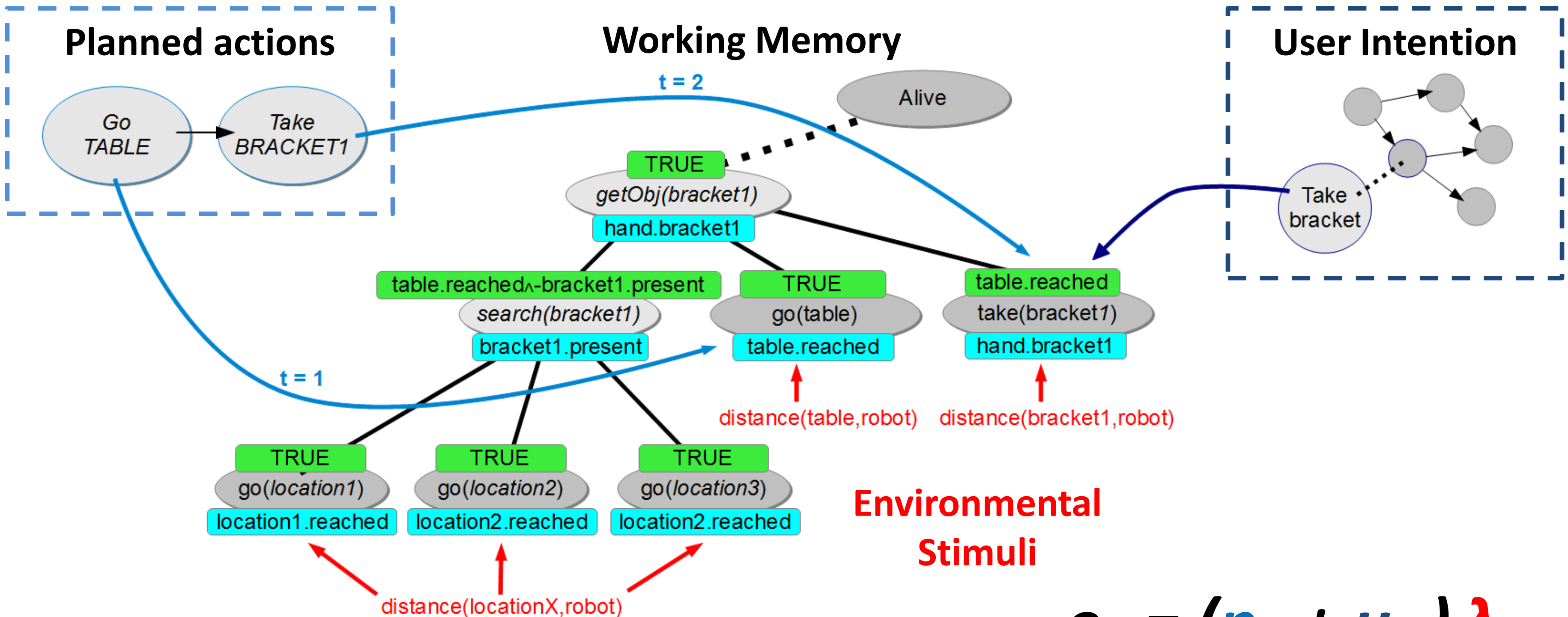
- The planner suggests to the attentional system actions or subtasks to perform.
- The attentional system integrates the plan with the user commands and the environmental context in order to flexibly react to unexpected events.

Hierarchical Task Network (HTN): Planning problems are specified in a hierarchical network providing a set of *primitive*, *compound* or *goal* tasks.

Human-Aware Task Planner (HATP): is an HTN planner that produces a different sequence of actions for each agent (human or robot).



Attentional Regulations



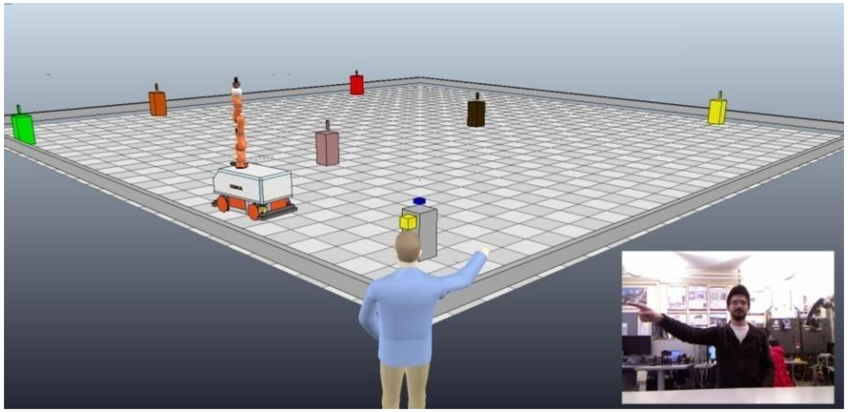
$$e_b = (p_b + u_b)\,\lambda_b$$

Attentional Regulation and Action Selection:

- The action selected by the plan and the user are **emphasized** (top-down).
- The bottom-up stimulations emphasizes actions that are **more accessible** to the robot.
- The **most emphasized** actions among all active actions are selected for the execution.

Testing and Results

Simulated Scenario: We consider the system at work in a simulated scenario where a mobile robot can execute pick-carry-and-place tasks in the presence of multiple colored objects.



PERFORMANCE			ERRORS	
Accuracy	**Precision**	**Recall**	**Violation**	**Worsening**
0.9439	0.9375	0.9868	0.8333	0.3333

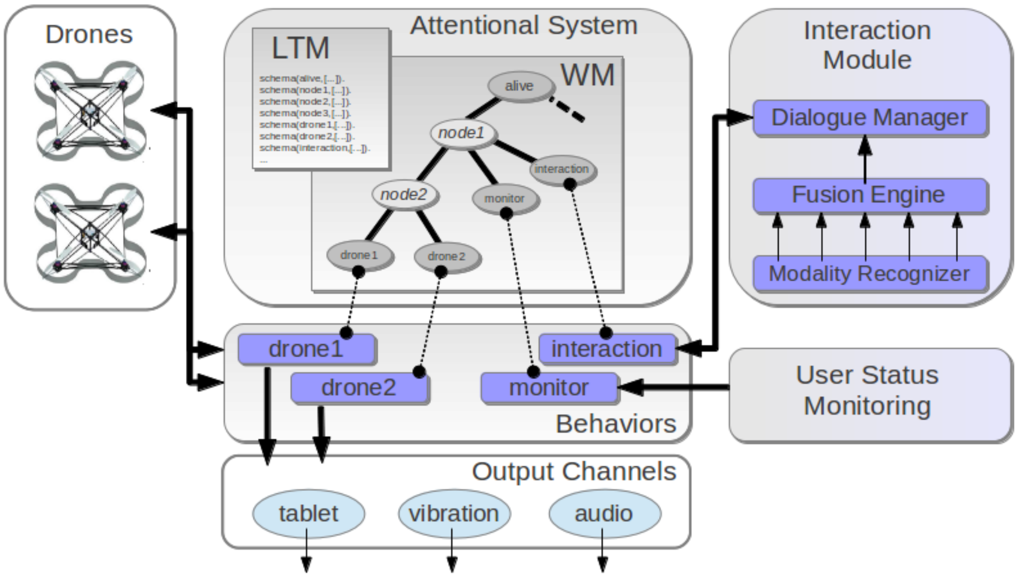
Real Scenario: We studied the system at work in a real-world scenario evaluating how the integrated system can flexibly adapt the execution to unexpected behaviors of the human avoiding replan.

The SHERPA Context

SHERPA

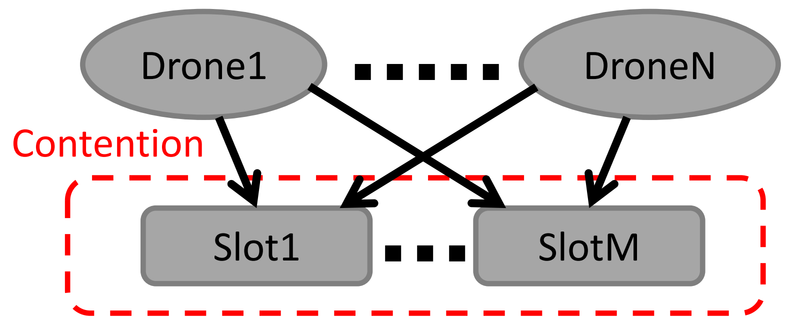
SEVENTH FRAMEWORK PROGRAMME

Attentional Cognitive Filtering



- Top-down and bottom-up attentional regulations used to emphasize relevant information.
- Hierarchical representation of the active tasks (WM).
- Monitoring processes with adaptive activations (frequency).
- Contention scheduling to manage conflicts (N-winner-take-all).

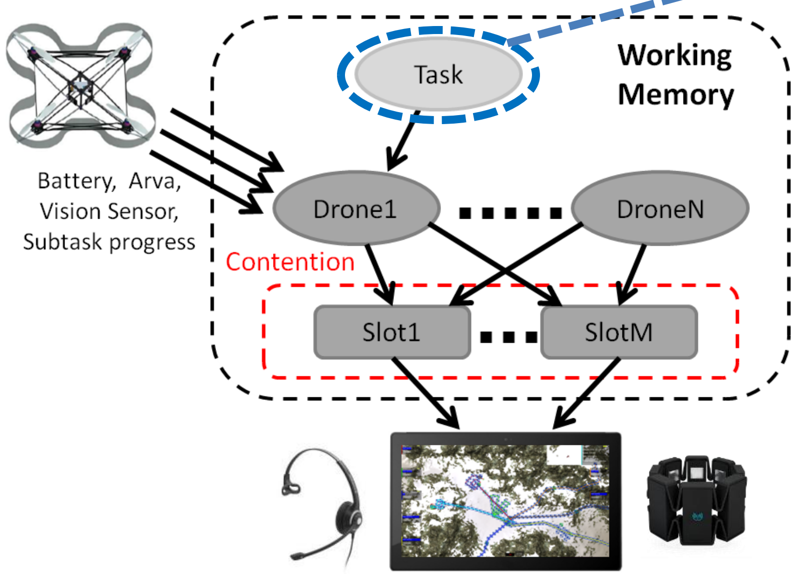
We assume the drones communication as a problem of contention for shared resources (communication slots).

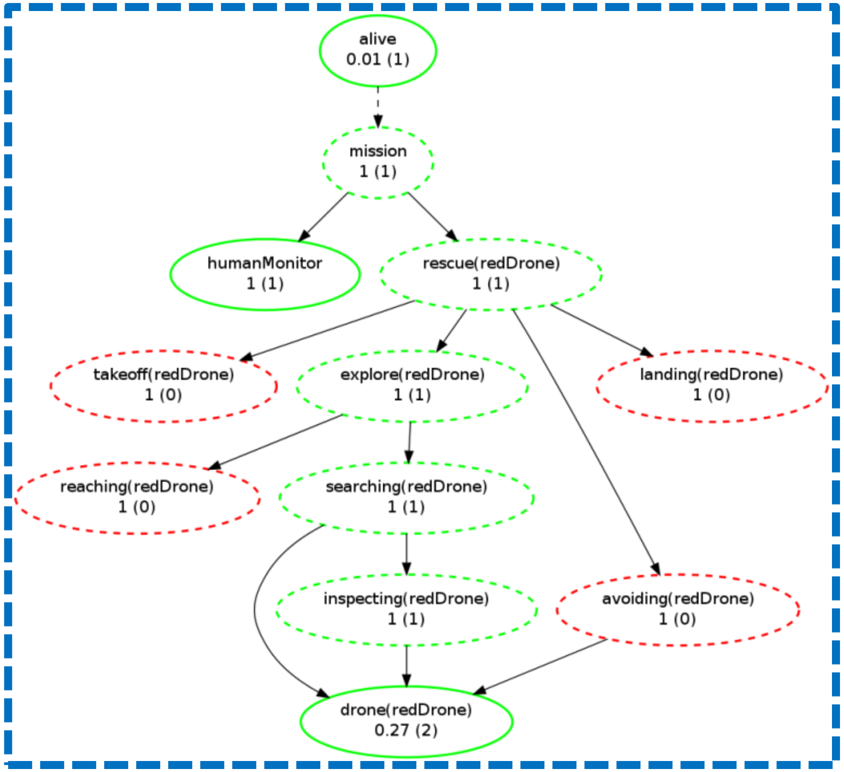


The number of available slots reduces with respect to the users cognitive load.

Attentional Regulations

- The **top-down** regulation is given by the mission state represented in WM by an annotated tree.
- The **bottom-up** regulation is given by the on-board sensors of the drones.

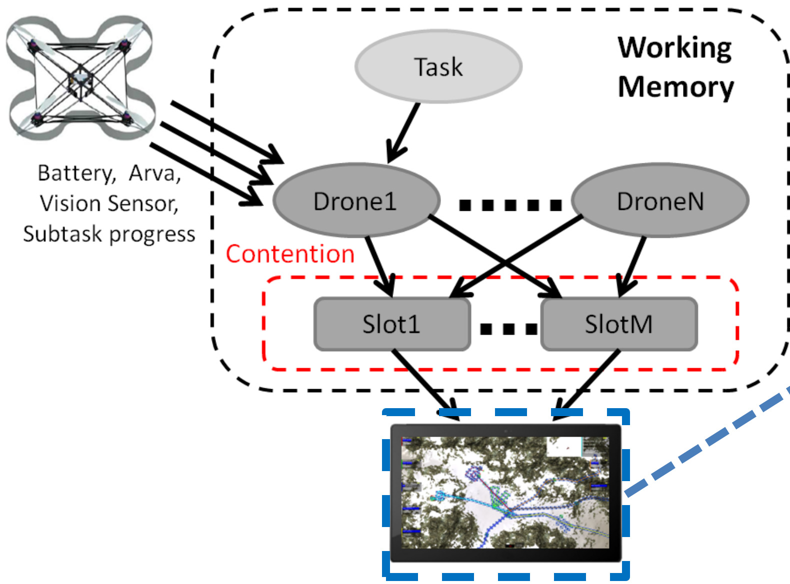


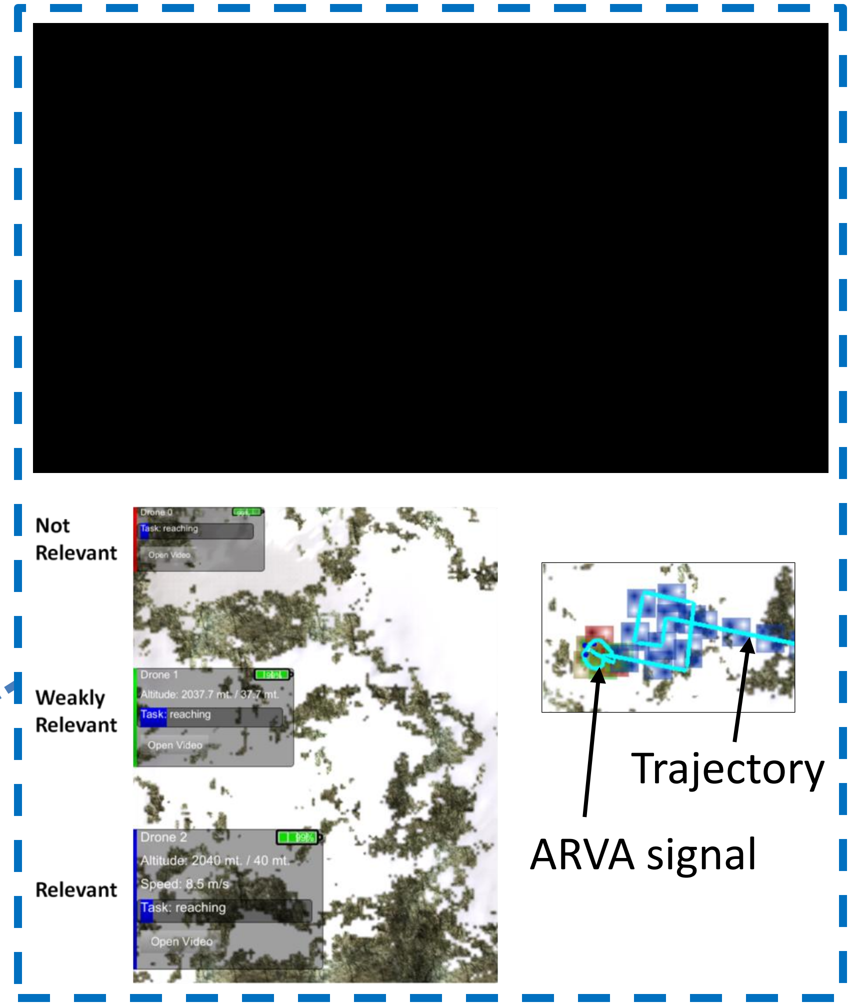


$$e = \mu_{task} (\lambda_{arva} + \lambda_{battery} + \lambda_{vision})$$

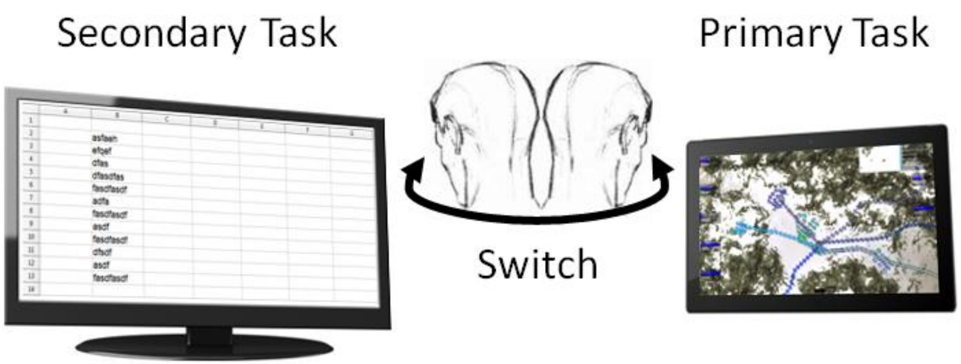
Adaptive Interface

- Background: map of the area with positions and trajectories
  - Multiple target monitoring [Alvarez & Cavanagh 2005]
- Foreground: info-box about the drones and streaming
  - Attentional shift [Posner et al. 1980]





Testing and Results



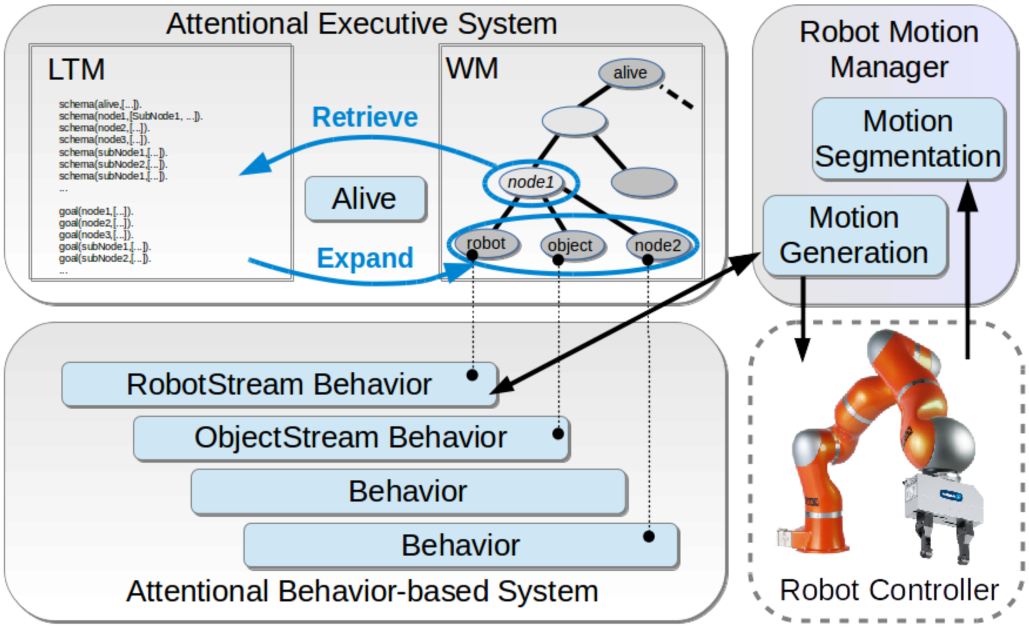
- **Interface-monitoring** (*primary*): the user has to monitor 8 drones during their search activities.
- **Word-counting** (*secondary*): the user should count the letters contained in each word of a list.

- **Adaptive**: the number of communication slots is reduced to one (maximum cognitive load).
- **No-adaptive**: fixed size of info-boxes and multiple parallel communications allowed.

	Events Occurred		Events Missed		Victims Missed	
	avg	std	avg	std	avg	std
No-Adapt	19.22	1.69	3.67	3	0.33	0.25
Adapt	11.44	2.28	1.22	0.69	0.11	0.11

no-Adaptive		Adaptive		Improvement	
avg	std	avg	std	avg	std
19.4%	1%	10.6%	0.4%	8.9%	0.4%

Structured Tasks Learning



- *The Robot Manager*: responsible for robotic motion, it splits the demonstration (teaching mode) in sub-segments.
- *The Attentional System*: contains the definition of all the **tasks available** to the system, updates the task according with the given sub-segments, select the most relevant action for the execution/learning.

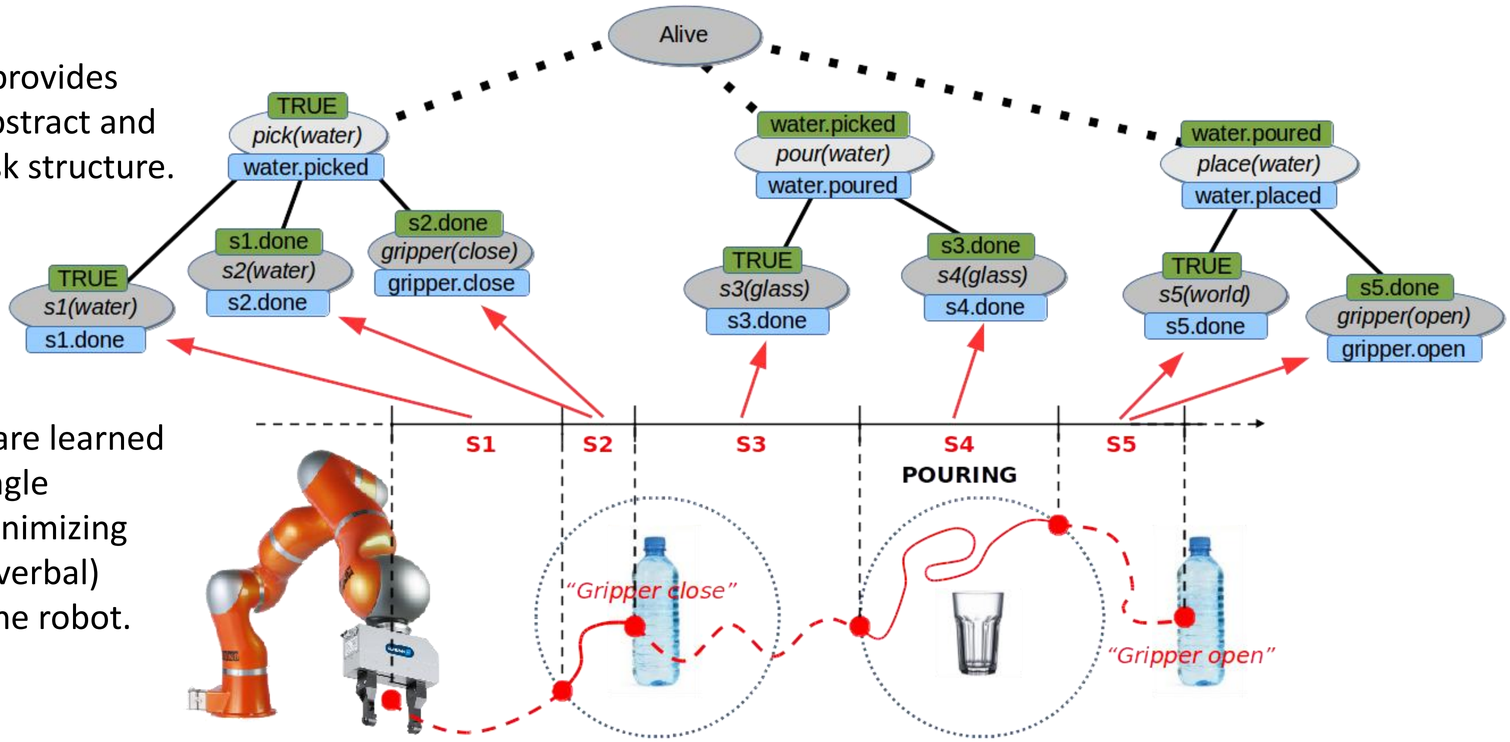
1. **Modality switching (teach/execute).**
   - *Kinesthetic Teaching.*
   - *Trajectory Execution.*
2. **Online actions segmentation.**
   - *Object Proximity:* end effector enters/leaves a near-object area.
   - *Open/Close Gripper:* open/close gripper command is received (speech).
3. **Online motion generation.**
   - *Far-Object-Action (FOA):* actions far from the object are executed with a point-to-point motion (less accurate).
   - *Near-Object-Action (NOA):* actions near the object are executed with Dynamic Movement Primitives (more accurate).

Segmentation and Task Update

- The attentional system tracks and monitors both the human and the robot task execution:
  - It is responsible for flexible and cooperative task execution;
  - During teaching it associates trained motion primitives to tasks and sub-tasks.

The user provides only an abstract and simple task structure.

Concrete actions are learned starting from a single demonstration minimizing the (physical and verbal) interaction with the robot.



Testing and Results

- **Add-water task** (*simple*): take and pour the water in the cup.
- **Prepare-coffee task** (*complex*): prepare the instant coffee.
- **Prepare-tea task** *(re-usage)*: prepare tea reusing some prepare-coffee subtasks.

We assess the system performances executing 10 learning sessions and 5 executions for each task (randomly changing objects positions) measuring times and success rate.

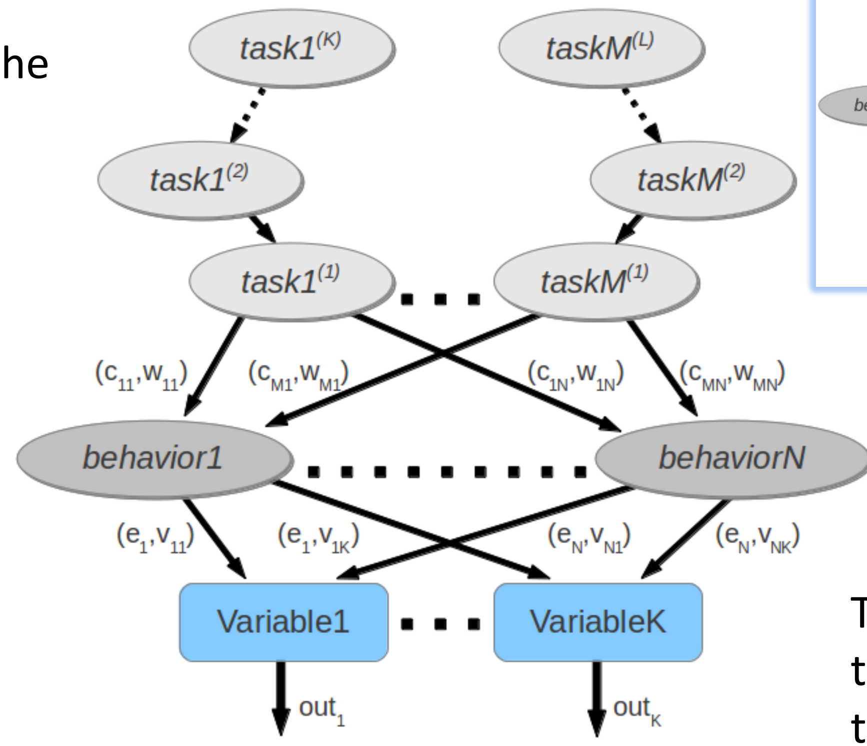
	Teaching time (avg ± std)	Execution time (avg ± std)	Success Rate
add-water	50.4 ± 2.0	77.5 ± 2.3	1
prepare-coffee	165.7 ± 12.8	224.5 ± 2.9	0.9

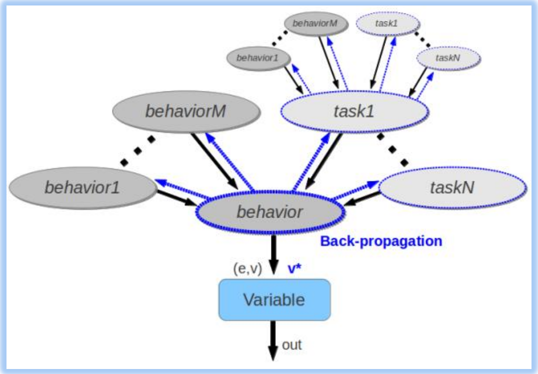
	Teaching time (avg ± std)
No Re-usage	84.2 ± 2.9
Re-usage	39.6 ± 2.4

47% faster!

Work in Progress…

Each edge of the WM can be associated with a weight that regulates the intensity of the attentional influence.





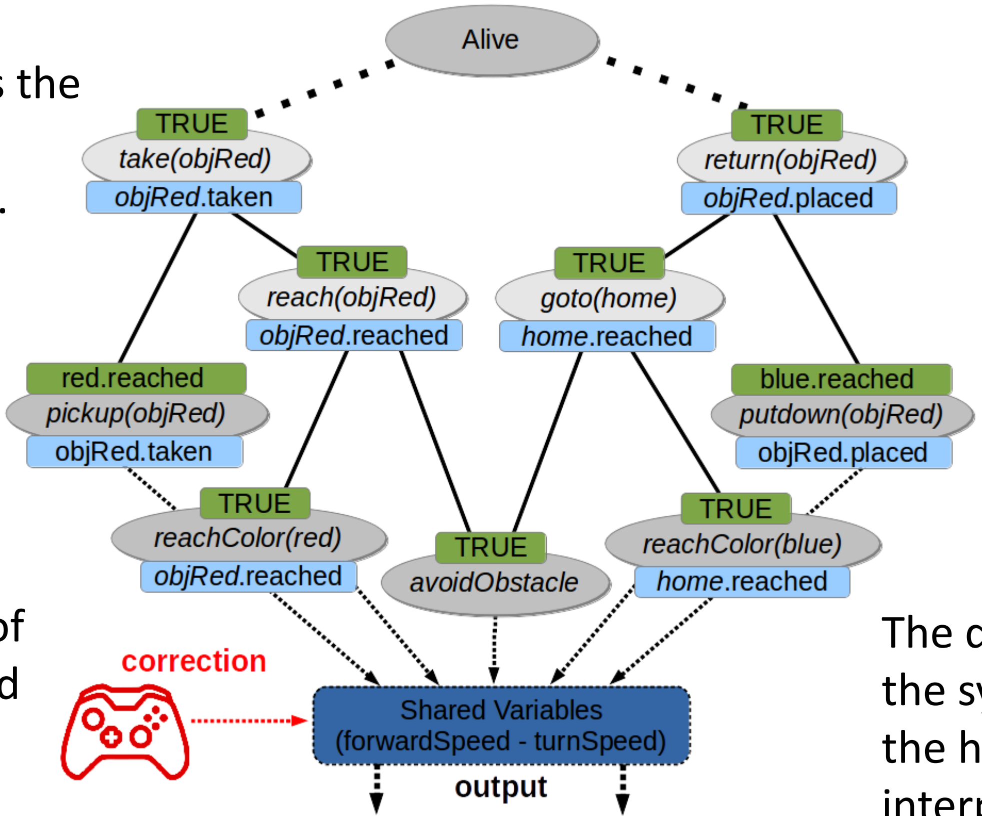
The *emphasis* value of each node is obtained as a weighted sum:

$$e_i = \Sigma w_{ji} c_{ji}$$

The difference between the system behavior and the human correction is interpreted as an error to be back-propagated.

Work in Progress…

Each edge of the WM can be associated with a weight that regulates the intensity of the attentional influence.



The *emphasis* value of each node is obtained as a weighted sum:

$$e_i = \Sigma w_{ji} c_{ji}$$

The difference between the system behavior and the human correction is interpreted as an error to be back-propagated.

Conclusions

- We proposed a robotic cognitive control framework for the flexible execution of structured complex tasks in the context of human-robot interaction.
- The framework was applied in different incremental scenarios, facing aspects of human-robot interaction, flexible plan execution, filtering and learning.
- Some of the applicative scenarios was also framed in the context of EU projects (SAPHARI and SHERPA).
- We assessed the framework performance in the different proposed contexts, considering simulations and real robotic scenarios.

Productions

Journal Papers:

- Flexible task execution and attentional regulations in human-robot interaction, IEEE Transactions on Cognitive and Developmental Systems, 2016.
- Kinesthetic teaching and attentional supervision of structured tasks in human-robot interaction, Autonomous Robots Journal, Springer, 2016. [Submitted]

Conference Papers:

- Attentional regulations in a situated human-robot dialogue. *IEEE RO-MAN*, 2014.
- Attentional top-down regulation and dialogue management in human-robot interaction. HRI, ACM, 2014.
- Plan Execution and Attentional Regulations for Flexible Human-Robot iteraction. IEEE SMC, 2015.
- Attentional supervision of human-robot collaborative plans. *IEEE RO-MAN*, 2016.
- Attentional multimodal interface for multidrone search in the Alps. IEEE SMC, 2016.
- Cognitive control and adaptive attentional regulations of robotic task execution. EuCognition meeting, 2016.

Book Chapters:

- Attentional Top-Down Regulations in a Situated Human-Robot Dialogue ICRA 2014 WS Robots in Homes and Industry: Where to Look First?
- Integrating Multimodal Interaction and Kinesthetic Teaching for Flexible Human-Robot Collaboration. HFR-2015.

Workshop Papers:

- Attentional Top-Down Regulations in a Situated Human-Robot Dialogue ICRA 2014 WS Robots in Homes and Industry: Where to Look First?
- Integrating Multimodal Interaction and Kinesthetic Teaching for Flexible Human-Robot Collaboration. HFR-2015.
- Attentional Plan Execution for Human-Robot Cooperation. AI*IA-2015 Italian Workshop on Artificial Intelligence and Robotics (AIRO2015).
- A human-robot interaction framework for search and rescue in the Alps, AI*IA-2016 Italian Workshop on Artificial Intelligence and Robotics (AIRO2016).
- Integrated task learning and kinesthetic teaching for human-robot cooperation, AI*IA-2016 Italian Workshop on Artificial Intelligence and Robotics (AIRO2016).

Thanks for Your attention