TAILOR

Foundations of Trustworthy AI – Integrating Reasoning, Learning and Optimization
TAILOR
Grant Agreement Number 952215

Foundations, techniques, algorithms and tools for allowing autonomous AI agents to decide and learn how to act - D5.2 Report

Document type (nature)	Report
Deliverable No	D5.2
Work package number(s)	5
Date	Due M42, February 29, 2024
Responsible Beneficiary	#6, UNIROMA
Author(s)	Giuseppe De Giacomo
Publicity level	Public
Short description	Report the novel insights, techniques, algorithms and tools developed within the scientific challenges tasks T5.1, T5.2, T5.3, T5.4

History			
Revision	**Date**	**Modification**	**Authors**
v1	2024-03-20	First version	Giuseppe De Giacomo

Document Review		
Reviewer	**Partner ID / Acronym**	**Date of report approval**
Luc De Raedt	KLeuven, #	2024-04-06
Fredrik Heintz	LiU, #1	2024-03-26

Table of Contents

Summary of the report

The main goal of WP5 is to promote the study of foundations, techniques, algorithms and tools for allowing autonomous AI agents to decide and learn how to act. The major challenge is integrating data-based methods with model-based methods by learning first-order symbolic models from non-symbolic data, to allow flexible and compositional reasoning and planning in terms of objects and relations. The interest in particular is to learn meaningful dynamic models from data that allow reasoning and explanation. Apart from the specific scientific work within the project itself, one of the most important objectives of this work-package is to pave the way for research on the topic of "How an AI agent decides and learns how to act" that is multidisciplinary, involving, planning, knowledge representation, synthesis and verification in formal methods, reinforcement learning in non-Markovian models, neuro-symbolic relational methods, and deep learning.

The WP is divided into four scientific challenge tasks, i.e., addressing four main scientific challenges in the theme, plus two extra tasks, one on cross-fertilization with industry and one on fostering a scientific community dedicated to this theme. Each scientific challenge task continuously interacts with the latter two in order to provide input and receive feedback and challenges in order to get a closed-loop approach to the research activities. The scientific challenges will have strong synergies with other WPs.

Introduction to the Deliverable

WP5 has two scientific deliverables D5.1 and D5.2. The first is an intermediate report (M24) and second one is the final report (M48, due at the end of the project). This document is D5.2.

Deliverables

D5.1: Foundations, techniques, algorithms and tools for allowing autonomous AI agents to decide and learn how to act v.1 (report, M24)

D5.2: Foundations, techniques, algorithms and tools for allowing autonomous AI agents to decide and learn how to act v.2 (report, M42)

Organisation

The following people have been involved in the Deliverable:

Partner Acronym	Name	Role
UNIROMA	Giuseppe De Giacomo	Leader of Task 5.1
Universitat Pompeu Fabra	Hector Geffner	Leader of Task 5.2
University of Basel	Malte Helmert	Leader of Task 5.3
Bruno Kessler Institute (FBK)	Paolo Traverso	Leader of Task 5.4
CNRS-IRIT	Andreas Herzig	Leader of Task 5.5
RWTH	Gerhard Lakemeyer	Leader of Task 5.6

Scientific Tasks

The work in WP5 explores the fundamental question: how does an AI agent decide and learn how to act? In particular, WP5 aims at empowering the agent to deliberate autonomously (i.e., without human intervention) how to act in the world. This objective strongly relates to the work in WP4, which focuses on learning how the world works and understanding its properties better.

WP5 aims at realising self-deliberating and autonomous systems by leveraging competencies in Planning and Knowledge Representation and Reasoning, as well as deep competencies in Learning and Optimization. Specifically, WP5 investigates issues like reasoning and planning for acting; learning strategies/plans from data; learning models from data and then do reasoning and planning; learning from past experiences and simulations for refining strategies/plans or models; monitoring the actual outcome of actions; recognizing possibly unexpected outcomes; reasoning, planning and learning how to deal with unexpected outcomes. In general, WP5 explores novel models of world dynamics and agent tasks, a new generation of solvers, and how to integrate data-based methods with model-based methods in deciding and learning how to act through several related scientific tasks. Crucially, empowering an AI agent with the ability to self-deliberate its behaviour and act autonomously carries significant safety risks, which must be guarded by human-guided specifications and oversight, meaning to find a balance of such power with safety. This aspect strongly relates to the agents' trustworthiness theme carried out in WP3.

The work in WP5 is divided into 4 “scientific challenge tasks”, i.e., addressing 4 main scientific challenges in the theme, namely:

- Task 5.1: Extended and multi-facet models of the world dynamics and tasks

- Task 5.2: Integrating data-based methods with model-based methods in deciding and learning how to act
- Task 5.3: Learning for reasoners and planners, and reasoners and planners for learning
- Task 5.4: Monitoring and controlling to make actions AI trustworthy in the real world

We describe the work in each of these challenges in the following. We must stress that most of the research done in this work package is of advanced foundational research in AI with a low technology readiness level (between TRL1 and TRL3). Though, we have a specific task on the potential impact of such foundational studies in future AI systems, also allowing for prospective cross-fertilization with industry to be carried out in WP8, namely:

- Task 5.5: Synergies Industry, Challenges, Roadmap concerning on autonomous actions in AI systems

Finally, we have a specific task dedicated to fostering a scientific community on this theme, namely:

- Task 5.6: Fostering the AI scientific community on the theme of deciding and learning how to act

We describe the work done in each of these tasks separately below. Each task continuously interacts with the others to provide input and receive feedback and challenges to get a closed-loop approach to the research activities. Moreover, synergies with the other scientific work packages WP3 (trustworthy AI), WP4 (Paradigms and Representation), WP6 (multi-agents and social), and WP7 (AutoAI) have been sought.

Task 5.1: Extended and Multi-facet Models of the World Dynamics and Tasks

The challenge of **deliberating the course of action** in partially known environments has been taken by AI mainly within Planning in nondeterministic domains. From the conceptual point of view, the essence of planning is program-synthesis under assumptions (assumptions being the model of the world). However, in Planning typically the task is simply to reach a desired state of affairs. Instead we want to consider agent tasks that can be sophisticated process specifications. For this reason, we adopt logical specification languages developed in Formal Methods. Specifically, we focus on Linear Temporal Logic (LTL), which is the specification formalism most used in Model Checking, except that, when we consider agent tasks, we adopt its finite trace variant LTLf. Observe that much of this work is related to what we may call “Assured Autonomy” i.e., an area of AI concerned with building AI agents that autonomously deliberate how to act under formal guarantees, which merges AI and Formal Methods (FM). Hence, much of the work done here is also related to WP3 (Trustworthy AI), which indeed includes studying AI systems with formal guarantees.

Beyond nondeterministic domains, partially known environments with multiple agents are the subject of **epistemic planning**, which is relevant for WP6 e.g. through its role in social interaction. The importance of the topic is witnessed by a special issue of the AI Journal with that title [Bolander et al. 2023]. Epistemic planning relies on the resources of epistemic logic in order to enrich the description of planning problems: initial state and goal are described in a language that is richer than that of classical planning, and the event models of dynamic epistemic logic cater for the description of epistemic action preconditions and effects. Results on planning with knowledge were published in the AI Journal [Cooper et al. 2021]. First results on the more difficult issue of planning with belief (which has to account for false beliefs and revision of beliefs) have been published at AAAI 2024 [Engesser et al. 2024]. Furthermore, strategic behaviour making use of epistemic reasoning was also investigated for the case of auctions [Mittelmann et al., 2021] and argumentation theory-based dialogues [Herzig and Yuste-Ginel, 2021c; 2021b; 2021a].

Focus on finite traces. The interest in finite traces comes from the consideration that, given a task, an intelligent agent should: (1) reason; (2) synthesise a course of actions; (3) execute such actions; and (4) be ready for the next task. If the task requires an infinite execution, then the agent would reason only once in its lifetime and then execute the synthesised program forever. While this is perfectly fine if the reasoning is done by the designer as in Formal Methods, it does not make much sense if the reasoning is done by the intelligent agent itself. This observation puts strong emphasis on finite traces in AI. The current advanced techniques for LTLf synthesis use symbolic encoding, with additional various forms of formula/automata decomposition. The very best technique available now has been developed in a paper at ICAPS 2021 [De Giacomo and Favorito, 2021]. We recently are looking at a radically different approach to LTLf synthesis based on forward search: from the formula one builds on the fly an AND-OR graph whose branching factor remains controlled through Knowledge Compilation techniques [De Giacomo et al., 2022c]. This approach shares some similarities with solvers for Fully Observable Non Deterministic (FOND) planning but has the capability of exploring a doubly exponential state space (as needed for LTLf synthesis) instead of a single exponential one. Based on such research, we recently developed two LTLf synthesis tools, respectively based on forward search techniques and on factorized symbolic encoding, see [Favorito, 2023] and [Favorito and Zhu, 2023]. Notably, these tools won, respectively, 1st and 2nd place in the LTLf track of the International Synthesis Competition SYNTHCOMP.

Pure Past LTL and Planning. An interesting observation arises when studying logic on finite traces. If the specification is given in Pure Past Linear Temporal Logic (PPLTL) then, because of a property of reverse automata, we can compute the corresponding Deterministic Finite-state Automaton (DFA) with at most one exponential blow-up [De Giacomo et al., 2020a]. As a result, one can build a symbolic representation of such a DFA in polynomial time in the size of the logical specification [De Giacomo et al., 2022a]. This observation is the base of techniques for planning in deterministic and nondeterministic domains for PPLTL formulas with only a polynomial overhead wrt standard planning. Specifically, we have shown that planning for PPLTL goals can be encoded into classical planning with minimal

overhead, introducing only a number of new fluents that is at most linear in the PPLTL goal and no spurious additional actions [Bonassi et al., 2023a] (the paper won the best student paper award at ICAPS 2023). We presented a similar encoding for planning in nondeterministic domains (FOND Planning). We showed that FOND planning for PPLTL goals can be polynomially translated into FOND planning with standard goals [Bonassi et al., 2023b]. Our empirical analyses confirm that such techniques perform better than existing ones for LTLf goals. Notably, the PPLTL-based techniques leverage existing state-of-the-art planning technologies, of which they essentially maintain the same scalability characteristics. For this reason, they have the potential to have a significant practical impact on the entire Planning in AI area. The use of PPLTL in the context of reinforcement learning is discussed later in the document.
Real-world applications often involve considering multiple agents. In the context of planning for temporally extended goals, an interesting future direction is to leverage reasoning techniques to multi-agent systems along the lines discussed in [Trapasso et al., 2023]. As different agents may have different tasks, adopting such techniques in multi-agent systems might first require agents to recognize goals of the other agents. In the context of temporally extended goals, this has been investigated in [Fraga Pereira et al., 2024].

Non-Markovian Environment Specification. While focusing on tasks specified by finite-trace formalisms, we cannot restrict environment specifications to finite traces. Indeed, the environment will not stop working when the agent finishes the task. This calls for finding a well-behaved way of representing the environment behavior. Among such formalisms are (nondeterministic) planning domains used in Reasoning about Actions and Planning. However, most of this research focused on Markovian behavior specification. That is, the next environment state is determined by the previous state, the agent action, and the environment response. Recently, this Markovianity has been challenged in several ways. We investigated synthesis (i.e., planning) for LTLf goals under LTL environment specifications in the case the agent must mandatorily stop at a certain point in [De Giacomo et al., 2021a] . Notably, LTL environment specifications capture (nondeterministic) planning domains as well as many interesting forms of non-Markovian properties, including fairness and stability. Of special interest are General Reactivity of Rank 1 (GR(1)) specifications, a form of safety that remains well-behaved computationally and maintains good scalability properties. We investigated synthesis for LTLf goals and GR(1) environment specifications in [De Giacomo et al., 2021b] and [De Giacomo et al., 2023]. We then turned the attention to forms of safety and guarantees properties (c.f. Manna&Pnueli's Safety-Progress LTL Hierarchy) as non-Markovian specification for both the agent task and the environment model. In such setting, we have devised synthesis techniques for arbitrary safety and guarantee properties that are still based on building a game arena out of the deterministic automata for the formulas (as for standard LTLf), but with more sophisticated game objectives [Aminof et al., 2023b].

Data-awareness. The work reported above does not consider unbounded data, which are needed for making our autonomous agent data aware. To do so we need to move from a propositional representation of the state to a first-order one. We can use Situation Calculus,

developed for Reasoning about Actions in AI as the main target framework for the lifting. In this context, we considered the following problems: (i) model checking and monitoring of first-order dynamic systems where new objects can be injected in the active domain during execution [Calvanese et al., 2022]; and (ii) Controller synthesis for manufacturing systems [De Giacomo et al., 2021d] [De Giacomo et al., 2022e]. We also extended Situation Calculus to deal with uncertainty and nondeterminism [De Giacomo and Lespérance, 2021]. In this context, we developed a general framework for abstracting the behavior of an agent that operates in a nondeterministic domain based on the nondeterministic situation calculus [Banihashemi et al., 2023]. We have also considered the case in which multiple agents act in the environment in the situation calculus. In this setting, we developed an approach for abstraction in multi-agent synchronous games [Lespérance et al., 2024].
Inspired by the Situation Calculus, we lately investigated extensions of classical planning with object creation, where problems are described in first-order logic, and allow for an unbounded number of objects. Specifically we show how one can leverage approaches for classical planning in the propositional setting by instantiating objects on the fly only when needed, getting novel techniques for first-order planning that inherits scalability from the propositional case [Correa et al., 2024].

One important aspect of data-awareness is handling data integration from multiple data sources. To do so, within the work for WP4 we further developed modern description logics, used for formalizing conceptual models and ontologies, see IJCAI 2021 [Console et al., 2021], JAIR 2021 [De Giacomo et al., 2021g], FI 2022 [Lembo et al. 2022].

Best-Effort and Multiple Environment Models. When an autonomous agent deliberates its course of actions, it may conclude that a strategy to fulfill the task regardless of the environment responses does not exist. When such a strategy does not exist, the agent can either give up or do its best. Along these lines, we introduced best-effort strategies in [Aminof et al., 2021b]. Based on the game-theoretic notion of dominance, best-effort strategies capture the rationality principle that a player should not use a strategy that is “dominated” by another of its strategies. As a result, best-effort strategies achieve the goal against a maximal set of environment responses and, interestingly, they always exist. Best-effort strategies share similarities and generalize so-called ``strong-cyclic plans” used often in FOND planning by taking advantage of stochasticity even if the goal cannot be reached, as we have shown in [Aminof et al., 2022]. Along these lines, we presented a synthesis technique for best-effort plans in FOND domains in [De Giacomo et al., 2023a]. However, the notion of best-effort strategy can be generalized to domains with non-Markovian supports and Borel goals, which generalize the classical planning in nondeterministic domains setting, as we discussed in [Aminof et al., 2023]. Leveraging the notion of dominance, we introduced dominant strategies in [Aminof et al., 2023a]. Dominant strategies achieve the goal against a maximum (vs. maximal) set of environment responses. An interesting observation is that best-effort and dominant strategies provide a suitable approach to agent reasoning with multiple environment models. For instance, an agent may be equipped with a nominal model of the environment and a model that also includes exceptional behaviors. An agent equipped with multiple environment models should be able to complete its task against all expected environment responses while acting reasonably to handle possible exceptional environment responses. Interestingly, best-effort strategies to do so always exist, as we have shown [Aminof et al., 2020]. This result extends to an arbitrary number of increasingly nondeterministic environment models, as we discussed in [Aminof et al., 2021a]. In practice, we have shown that existing implementations of LTLf best-effort synthesis techniques are quite promising, exhibiting only a minor overhead compared to standard existing LTLf synthesis techniques [De Giacomo et al., 2023b].

Learning and reasoning. Our focus on non-Markovian/Temporal tasks and dynamics leads towards non-traditional forms of reinforcement learning, which merge reasoning and learning. We have developed the notion of Restraining Bolts as an LTLf specification that limits and controls the reinforcement learning process, steering it toward creating strategies to satisfy as much as possible the LTLf specification [De Giacomo et al., 2019]. We have investigated learning the restraining bolt itself through imitation learning [De Giacomo et al., 2020b]. We have considered forms of restraining bolt specifications that are stemming out of concepts developed in monitoring i.e., defining rewards that depend on prefixes of the desired traces instead of the whole trace itself. More recently we turned to approaches for ``shielding” reinforcement learning to avoid that the synthesized policy breaks certain safety constraints, both during the training phase and the execution case. Specifically, taking advantage of the special characteristic of PPLTL specifications discussed above, we studied approaches for safe reinforcement learning via action masking with PPLTL specifications, developed into a framework, Pure-Past Action Masking (PPAM), which inherits from the use of PPLTL for specifying safety, excellent scalability properties [Varricchione et al. 2024] (this research was initiated by a visit by a research group of Utrecht University to Sapienza, thanks to the TAILOR’s Connectivity Funds).

Another line of research is that of non-Markovian reinforcement learning in which the agent learns non-Markovian dynamics and rewards, still implicitly based on LTLf specifications,developing solutions that simultaneously learn a DFA and solve related Markov Decision Problems (MDPs), see [Ronca and De Giacomo, 2021b] and IJCAI 2022 [Ronca et al., 2022b]. Along these lines, we studied automata cascades, decomposition of finite state automata into a set of prime automata [Ronca et al., 2023]. (c.f. Prime Decomposition Theorem from Krohn and Rhodes). Automata cascades open to the possibility of learning automata representing large dynamical systems consisting of many interacting components. As the applicability of reinforcement learning techniques is limited by the large number of required samples, we considered a linear hierarchy of abstraction layers of MDPs, i.e., each layer is an MDP representing a coarser model of the one below. Within this framework, we investigated a novel form of Reward Shaping that guides the learning process towards the most complex domain [Cipollone et al., 2023b].

We investigated the use of inductive logic programming for learning LTLf formulas from sets of example traces [Ielo et al., 2023] (the paper won the best paper award at the ICL2023 conference). In particular, the techniques developed in this work are of interest in Business Process Management and Smart Manufacturing (more details below).

In the context of neuro-symbolic approaches to ML, we investigated symbol grounding, i.e., mapping high-dimensional raw data into an interpretation over a finite set of abstract concepts with a known meaning, without direct supervision, using symbolic knowledge in the form of LTLf specifications [Umili et al., 2023].

Business Process Management (BPM) and Industry 4.0 Smart Manufacturing. Within BPM, DFAs are largely employed as foundation mechanisms to perform formal reasoning tasks over the information contained in the event logs. Given the strong connection among temporal specifications over finite traces and DFAs, BPM is a promising field of application of LTLf and PPLTL reasoning and learning techniques. Many business processes (BP) can indeed be declaratively specified by suitable LTLf or PPLTL formulas. In fact, variants of LTLf have already been used to describe possible execution traces of BPs [Chiariello et al., 2023]. Among our applications there are: (i) monitoring, i.e., the task of checking that the execution of a process is compliant with its declarative specification [De Giacomo et al., 2022f]; (ii) trace alignment, the task of aligning real process execution traces to the declarative specification of the process [De Giacomo et al., 2021f], which shares many similarities with planning [De Giacomo et al., 2023c]; and (iii) controller synthesis, the task of synthesising plans that delegate tasks in a supplied process recipe to the available manufacturing resources [De Giacomo et al., 2022e]. Smart manufacturing processes can also be expressed in terms of precedence relations between manufacturing activities that can be captured by suitable LTLf formulas as well. As a result, combining our recently developed smart manufacturing techniques, such as Digital Twins and API orchestration via Markov Decision Processes, see [De Giacomo et al., 2023e] and [De Giacomo et al., 2023d], respectively, and LTLf reasoning and learning techniques, is promising to improve the autonomy and resilience of industry 4.0 smart systems. All techniques presented above can also be leveraged with the recently developed approaches that learn finite-state automata or LTLf specifications from input data, see [Agostinelli et al., 2023] and [Ielo et al., 2023], respectively.

Other directions. In many scenarios, autonomous agents do not have to commit to some specific strategy. Instead, they must be able to adjust their course of actions if circumstances require them. An important result from Discrete Event Control says that there exists a “maximally permissive nondeterministic strategy” when we focus on safety specifications i.e., one which allows the agent to autonomously decide the next action to perform. Unfortunately, for LTLf task specifications this result does not apply. Nonetheless, the “maximally permissive nondeterministic strategy” can be constructed also in the setting of intelligent agents. The construction shown in [Zhu and De Giacomo, 2022b] shows how to do it by suitably composing two nondeterministic strategies. Along the same lines, we have investigated strategies that allow the agent to adapt its course of actions to achieve different forms of goals, called “duties” and “rights”, during execution [Zhu and De Giacomo, 2022a]. Related to multiple environment models and imitation learning is the problem of mimicking behaviours in separated domains. That is, the task of synthesising a strategy that, step-by-step, maps every behaviour in some dynamic domain DA into a behaviour in another dynamic domain DB so that a certain input specification is met. This problem is of interest in, e.g., human-robot cooperative tasks where robotic agents must mimic the behaviour of a human operator, as we discussed in [De Giacomo et al., 2023f]. Recently, we also investigated how to leverage the expressivity of LTL by using second-order quantifiers and adopting a “behavioural” semantics. The resulting logic, called Behavioral QLTL, can naturally capture planning in nondeterministic domains and LTL synthesis tasks through formulas with a simple quantification alternation [De Giacomo and Perelli, 2023].

Task 5.2: Integrating Data-based Methods with Model-based Methods in Deciding and Learning How to Act

Task 5.2 is about the use of data-based methods with model-based methods in deciding and learning how to act. This task was led by Hector Geffner until January 2023, and since then by Anders Jonsson, both at UPF. The aim is to study the foundations, techniques, algorithms and tools for integrating data-based learning methods with model-based methods for acting and planning. The integration of learning and reasoning (planning) methods is critical in AI, where current (deep) learning-based methods deliver reactive and opaque boxes ("System 1") that do not generalise properly and make no attempt to understand their environments, while model-based methods rely on models that must be supplied by hand in a suitable language. A key challenge is thus learning symbolic model representations automatically from data, as well as representations of general policies and general problem structure.

We have thus addressed these challenges on three main fronts: learning compact and general representations of environment dynamics (action models or simply models), learning representation of action strategies that generalise to classes of "similar" problems, and learning the subgoal structure of these potentially infinite classes of "similar" problems.

Model learning (Action models, Dynamics). One of the advantages of traditional model-based methods in AI planning is that it provides a crisp notion of "problem similarity" that follows from the languages used for modelling planning problems over discrete state spaces. In fact, two planning problems are deemed "similar" when they are instances of the same planning domain. A planning domain is given by a set of action schemas with lifted preconditions and effects defined in terms of a fixed set of domain predicates. The common structure of a domain given by the action schemas and predicates captures precisely what is common among all "similar" problems, and it is key for obtaining policies and problem subgoals that generalise across all "similar" problems.

A key question is how these first-order symbolic planning domains, captured by action schemas and predicates, can be learned from non-symbolic data. In works reported in ECAI 2020 [B. Bonet, H. Geffner] and in KR 2021 [Rodriguez et al., 2021a], we show how such planning domains can be learned from the state graphs reflecting the structure of the state space of small instances alone. It is well known that a planning instance determines a unique state graph; these approaches address the inverse problem: learning the simplest planning instances over a common domain, that is not known, that generates the observed state graphs. This problem is cast and solved as a combinatorial optimization problem expressed either as a Weighted-MAX SAT problem or as an answer set program. In these works, the states are deemed as "black boxes". More recent work reported at the 2022

ICAPS Workshop on Bridging the Gap Between AI Planning and Reinforcement Learning, considers states expressed as 2-dimensional parsed images, and a similar approach is used to learn planning representations that are grounded on such parsed images [A. Occhipinti, B. Bonet, H. Geffner, 2022].

Another alternative is to learn symbolic model dynamics in the form of finite state automata. In such a representation, the automaton state captures the history of agent interaction, which is necessary to determine the best future course of action. In work published in ICML [Furelos-Blanco et al., 2023], we show how to learn finite state automata in the form of reward machines from traces of high-level events. An agent can then combine the reward machines with reinforcement learning in order to learn a near-optimal policy. In recent work published at NeurIPS [Cipollone et al., 2023], we relax the assumption that the agent has prior knowledge of high-level events, and learn finite state automata directly from histories of raw action-observation pairs.

Learning General Policies. Even with a known compact model, planning is (NP) hard. It is thus natural to ask how to make planning simpler by suitable preprocessing, exploiting the knowledge of a planning domain. In the extreme case, one wants to learn general policies over all possible instances of a given or learned domain. Such policies are called general policies as they are not tied to a particular planning problem nor to a particular state space. A lot of the work in deep learning and deep reinforcement learning has been devoted to learning such policies, in most cases, without learning or using first-order domain representations. The results then have not been robust, nor transparent. In the context of TAILOR, we have thus looked at the problem of learning general policies by leveraging both data and first-order domain models (given or learned, as above). For this, it is crucial to define a domain-independent language for expressing general policies and a way of learning such policies from the data and the models. In work reported at AAAI 2021 [Francès et al., 2021], a novel formulation for expressing and learning such policies is developed. Moreover, the learned general policies are proved to be correct as well. Another paper at AAAI 2021 established correspondences between general policies and the notion of planning width [Bonet and Geffner, 2021], while the works at ICAPS 2022 and KR 2022 have shown how similar policies can be learned using graph neural networks [Ståhlberg et al., 2022a,b]. The ICAPS 2022 work was distinguished as the conference Best Paper. The work at KR 2023 shows how general policies represented as graph neural networks can be effectively learned using policy gradient methods from RL [Ståhlberg et al., 2023].

Another line of research is to compute general policies in the form of planning programs that can be used to solve any instance in a given family of sequential decision problems. In work presented at ICAPS [Segovia-Aguas et al., 2021] and SoCS [Segovia-Aguas et al., 2022a], we show how to compute planning programs using a pointer-based formulation that decouples the structure of the planning program from the actual objects of a specific instance. Instead of applying actions on specific objects, our formalism applies actions on pointers, and the outcome of the action depends on which objects are currently referenced by the pointers. The pointer-based representation makes it possible to scale to much larger instances and compute planning programs that generalise better to any instance. This work has recently been published in a journal version [Segovia-Aguas et al., 2024].

Learning Subgoal Structure (Sketches). The problem of learning subgoal structure appears in both planning and in reinforcement learning where learning in the presence of sparse rewards has been particularly challenging. Some methods developed in the RL setting for addressing this problem involve extra subgoal information in the form of intrinsic rewards, reward machines, or restraining bolts. In most cases, this extra information is provided by hand, in some cases, it is learned automatically, although there has not been a good theory of what the subgoals of a problem should be. Recently, we have developed both a language and a theory for expressing and characterising subgoal structure in the form of sketches. The language of sketches is similar to the language of general policies but the semantics is slightly different. Roughly, a general policy implicitly defines which state transitions (s,s') are "good". The same idea is used in sketches but the state s' does not have to be 1-step away from s; it can be a possible subgoal from s. A sketch thus decomposes problems into subproblems and a sketch is "good" if the resulting subproblems have all bounded widths, and hence can be solved in polynomial time. A paper at AAAI 2021 introduced the language and semantics of sketches [Bonet and Geffner, 2021], and follow up work at KR 2021, showed how to use sketches for taming the complexity of known benchmark domains in planning [Drexler et al., 2021], and a more recent work at ICAPS 2022, how to learn sketches automatically given the common planning domain and some small domain instances [Drexler et al., 2022].

A model and learning algorithm for hierarchical sketches is presented at KR 2023 [Drexler et al., 2023]. A recent paper on how to exploit sketches for classical planning was accepted at ICAPS 2024 [Drexler et al., 2024].

The language of sketches is extended with internal states, registers, and wrapped into modules that can call other modules, into a more expressive and modular language for problem decomposition; it was first presented at the GenPlan workshop at NeurIPS 2023 [Bonet et al., 2023], and then accepted at ICAPS 2024 [Bonet et al., 2024].

Exploiting subgoal structure is precisely the topic of hierarchical reinforcement learning. In work presented at CoLLAs [Steccanella et al., 2023], we learn abstract representations which can be used to define subtasks for hierarchical decomposition. One known limitation of hierarchical reinforcement learning is the tradeoff between optimality and computational efficiency. In work published at EWRL [Infante et al., 2023] and accepted at ICAPS [Kuric et al., 2024], we show for the first time how to achieve global optimality in hierarchical reinforcement learning, thus exploiting the benefits of hierarchical decomposition without sacrificing the ability to select globally optimal actions. In related work published at AISTATS [Bourel et al., 2023], we prove that hierarchical decomposition in the form of reward machines leads to lower regret than what is possible using a flat representation.

Other. Other relevant works published in this period include a paper at [Bonet and Geffner, JAIR 2020] detailing a planning model, called Qualitative Numerical Planning or QNP, that is suitable for expressing the model abstractions that general policies are aimed to solve [B. Bonet, H. Geffner, 2020] and a paper at ICAPS 2021 [Rodriguez et al., 2021b], developing general algorithms for solving both QNPs and fully-observable, non-deterministic (FOND) planning problems. This work was distinguished as Best Paper at ICAPS 2021. Likewise, our work on learning first-order formulas for characterising dead-end states in planning obtained a Distinguished Paper award at IJCAI 2021 [Stahlberg et al., 2021]. Finally, an overview of our work on learning representations for acting and planning appeared at AAAI 2022 [Geffner, 2022].

Task 5.3: Learning for Reasoners and Planners, and Reasoners and Planners for Learning

Task 5.3 is concerned with the integration of data-based machine learning methods with model-based planning techniques with the aim of developing and studying foundations, techniques, algorithms and tools for integrating learning into reasoners and planners. The challenge is to overcome the fundamental differences of the two approaches: On the one hand, learning approaches are not based on a given model. They are therefore inductive and tend to come without guarantees on correctness or optimality, but the last decade has shown that learning can advance the state of the art significantly in a wide variety of applications. Reasoning approaches, on the other hand, exploit a given model. They deduce a solution from formal (often logic-based) representations and yield a solution with guaranteed soundness or optimality guarantees. The drawback is that reasoning approaches can be somewhat rigid and inflexible, lacking ability to adapt to specific applications.

There are different ways that allow us to exploit the potential of learning techniques in a reasoning system without sacrificing correctness and optimality guarantees. Using a provided model as a black-box simulator for a learning algorithm has become popular with the development of Monte-Carlo Tree Search [Kocsis & Szepesvári, 2006] and opened the path for well-informed heuristic and policy networks that guide the search performed by such algorithms [Silver et al., 2016].

A combination of reasoning and learning that maintains the advantages of both is to use learned information in the deduction process of reasoning processes.
We focus here on different ways to combine reasoning and learning that maintain the advantages of both by using learned information in the deduction process of a reasoning process. Progress that exploits this idea has been achieved by learning heuristics for planning, learning to decompose planning tasks into subtasks, dynamic algorithm configuration and planner selection, as well as using learning for plan recognition.

Neural Network Heuristics for Planning Heuristic state space search is among the most successful approaches for automated planning, and improvements with respect to the heuristic (e.g., improved quality or computation time) have been responsible for many advances of the state of art. Heuristic functions are therefore an obvious starting point to introduce machine learning techniques into planning, in particular since planning with a learned heuristic remains sound as long as mild constraints on the safety of the heuristic hold.

Neural networks (NN) are commonly used to learn a function based on labelled input data. In the case of planning heuristics, NNs are trained to approximate a function that takes a state as input and outputs a distance estimate for the state. There are different strategies for the generation of training data, including the computation of shortest paths. As this is only possible in small instances, for states that are sufficiently close to the goal or in domains where data in small instances generalises well to larger ones, [Ferber et al., 2022a] introduce a procedure that makes training data generation scalable through bootstrapping and approximate value iteration, and [Micheli and Valentini, 2021] mitigate this issue by using deep reinforcement learning to train a NN-based heuristic.

Recent NN-based function approximation techniques provide not just the approximated function estimate, but also information on how confident the estimator is in its estimate. In an application like search with NN-based heuristics, it is a natural idea to take this into account, e.g. by falling back to a standard heuristic where confidence is low. [Heller et al., 2022] explore this idea by performing search with multiple open lists that depend on the confidence of heuristic values.

Learning State Space Structure Several works within Task 5.3 investigated how to learn and exploit the state space structure of a planning task to improve the performance. In their analysis of Greedy Best-First Search (GFBS), Heusner, Keller and Helmert [IJCAI 2018, pp. 5269-5273] discovered that there are states in a state space topology that can be used to partition each run of GBFS into phases. Unfortunately, these progress states can only be determined after the search has terminated successfully, and they have therefore been of limited practical interest. Combining learning of states with generalisation among instances in the same domain allows to determine progress states on small instances, compute a formula over description logic features that generalises well over all instances of the domain and exploit the generalised information to scale to large instances. [Ferber et al., 2022b] take a first step in this direction by implementing the sketched algorithm. Their work exploits the learned information by breaking ties in favour of progress states.

Recent works considered landmarks – properties that must be satisfied by all plans – as a base component of cost-optimal solution planners. Landmark-based heuristics exploit that each landmark must be achieved at least once. However, if the orderings between the landmarks induce cyclic-dependencies, one of the landmarks in each cycle must be achieved an additional time. Recently, [Büchner et al., 2021] introduced novel heuristics for cost-optimal planning that consider cyclic-dependencies between landmarks in addition to the cost for achieving all landmarks once. They show that their heuristics dominate the minimum hitting set solution over any set of landmarks as well as h^+ if all delete-relaxation landmarks are considered. It should be noted that [Büchner et al., 2021] has been nominated as Runner-Up for Best Student Paper Award at ICAPS 2021. A further application of landmarks in heuristic search is landmarks progression. In this setting, [Büchner et al., 2023] presented a general framework for using landmarks in any kind of best-first search. Its core component, the progression function, uses orderings and search history to determine which landmarks must still be achieved. They defined a sound progression function that allows to exploit reasonable orderings in cost-optimal planning and show empirically that their new progression function is beneficial both in satisficing and optimal planning. Notably, [Büchner et al., 2023] won the Best Paper Award at ICAPS 2023.

Interesting alternative heuristic approaches are pattern databases (PDBs). [Sievers et al., 2022] have shown how to compute pattern database (PDB) heuristics for decoupled states in decoupled search, a planning approach where variables are factored into a center and several leaf factors. They show that, in the general case, for arbitrary collections of PDBs, computing the heuristic for a decoupled state is exponential in the number of leaf components of decoupled search. However, they derive several variants of decoupled PDB heuristics that allow to additively combine PDBs avoiding this blow-up and evaluate them empirically. Their experimental study suggests that PDBs can yield strong performance in decoupled search, surpassing the previous state-of-the art in optimal decoupled search. Notably, [Sievers et al., 2022] won the Best Paper Award at SoCS 2022.

While many heuristic planning approaches are concerned with exploiting the state space structure to improve the search performance, grounding planning tasks, i.e., translate planning tasks specified in a first-order language (e.g. PDDL) has become a larger bottleneck. Building on the most common grounder, which uses Datalog to find all reachable atoms and actions, [Correa et al., 2023] investigated a new method to ground planning tasks which uses tree decompositions and iterated solving. Their algorithm can ground more instances than the baseline, and most tasks it cannot ground are out of reach from any ground planner. Notably, [Correa et al., 2023] was nominated as Runner-Up for Best Student Paper Award at ICAPS 2022.

Yet another method to improve the performance of classical planners is cost partitioning, a method which allows combining multiple admissible heuristics while retaining an admissible bound. In their recent work, [Kloßner et al., 2022] investigated how to extend cost partitioning in classical planning to probabilistic planning by generalising deterministic transition systems to stochastic shortest path problems (SSPs). Specifically, they show that cost partitioning still holds in their extended theory and how to optimally partition costs for a large class of abstraction heuristics for SSPs. Finally, they analyse occupation measure heuristics for SSPs as well as the theory of approximate linear programming for reward-oriented Markov decision processes. Notably, [Kloßner et al., 2022] was nominated as Runner-Up for Best Student Paper Award at ICAPS 2023.

Dynamic Algorithm Configuration and Selection Another idea pursued in the context of task 5.3 is to use machine learning to decide which algorithm to apply in which situation or how to configure an algorithm in dependence of the current situation. There are different parts of a heuristic search algorithm that can be selected based on this idea, and in the extreme case the entire planning algorithm is determined dynamically before search has even started.

There have been different ideas that select the right planner based on features of tasks, and the most successful ones are based on NNs [Sievers et al., AAAI 2019; Ma et al., AAAI 2020]. The drawback of the neural network approaches is that the learned models are not interpretable, i.e., it is not clear why a planner is selected and which task features are actually important for the selection. [Ferber and Seipp, 2022] show that complex black-box models are not required to learn strong planner selectors. They train a decision tree which yields equally strong results and allows insights why and in which situation certain planners are selected.

Dynamic algorithm configuration [Biedenkapp et al., ECAI 2020] is a meta-algorithmic approach that uses information about the internal behaviour of an algorithm and information about the instance it is run on to change the configuration of the algorithm during its execution. [Biedenkapp et al., 2022] apply the idea to learn an open list selection policy in a principled and data-driven manner. They not only show that it is possible to learn strong selectors but also gain insights on the advantage of the manually generated open list selection strategy of the LAMA planner [Richter and Westphal, JAIR 2010] .

Learning in Plan Recognition Plan recognition is the task of inferring the actual plan an observed agent (possibly with noisy observations) is performing to achieve a goal. [de A. Santos et al., 2021] and [Rosa Amado et al., 2021] follow two different approaches to tackle the issue: The former encode the problem as a linear program and use reasoning alone to tackle the problem, whereas the latter develop a novel approach to solve both goal and plan recognition tasks simultaneously by combining planning and machine learning techniques to mitigate problems of low and faulty observability. A set of plans is used to train a predictive statistical model of the most likely next states given a set of state observations, and combining these predictive models with landmark heuristics allows to predict the most likely next state given a sequence of observations.

Task 5.4: Monitoring and Controlling to Make Actions AI Trustworthy in the Real World

The main objective of this task is to study foundations, techniques, algorithms and tools for devising and learning meaningful dynamic models that mix human understandable fluents versus hidden data features. In particular, we are interested in updating and correcting imperfect models, detecting problems in a model; learning from failures; learning (soft) constraints on the model when the model fails; mixing prior human dynamic knowledge/models with learning from data.

We worked on aspects concerning the goal of mixing prior human dynamic knowledge [Lamanna et al., 2022] proposed a method where an agent can exploit its symbolic knowledge about a planning domain by grounding it in the environment in which it operates. When the environment is initially unknown by the agent, the agent needs to explore it and discover the salient aspects of the environment necessary to reach its goals. Namely, the agent has to discover different aspects of the environments in which it operates. In particular it has to discover the objects present in the environment, the properties of these objects, and their relations, and finally how abstract actions can be successfully executed. We devise a framework that aims to accomplish the aforementioned perspective for an agent that perceives the environment partially and subjectively, through real value sensors (e.g., GPS, and on-board camera) and can operate in the environment through low level actuators (e.g., move forward of 20 cm). We evaluate the proposed architecture in photo-realistic simulated environments, where the sensors are RGB-D on-board camera, GPS and compass, and low level actions include movements, grasping/releasing objects, and manipulating objects. The agent is placed in an unknown environment and asked to find objects of a certain type, place an object on top of another, close or open an object of a certain type. We compare our approach with a state of the art method on object goal navigation based on reinforcement learning, showing better performances. This work builds on top of a general environment for acting, learning, and planning that has been developed in [Lamanna et al., 2021a, Lamanna et al., 2021b].

At the end of 2022 and 2023 we started tackling the problem of planning for learning in an unknown environment. In [Lamanna et al. 2023a] we present an online method based on symbolic planning that allows an agent embedded in a physical environment to recognize objects and their properties from sensory data. Such a perceptual ability is often implemented by supervised machine learning models, which are pre-trained using a set of labelled data. In real-world, open-ended deployments, however, it is unrealistic to assume to have a pre-trained model for all possible environments. Therefore, agents need to dynamically learn/adapt/extend their perceptual abilities online, in an autonomous way, by exploring and interacting with the environment where they operate. This paper describes a way to do so, by exploiting symbolic planning. Specifically, we formalise the problem of automatically training a neural network to recognize object properties as a symbolic planning problem (using PDDL). We use planning techniques to produce a strategy for automating the training dataset creation and the learning process. Finally, we provide an experimental evaluation in both a simulated and a real environment, which shows that the proposed approach is able to successfully learn how to recognize new object properties. In [Lamanna23b] we developed an approach that enables autonomous agents embedded in a physical environment to correctly perceive the state of the environment from sensory data. In partially observable environments, certain properties can be perceived only in specific situations and from certain viewpoints that can be reached by the agent by planning and executing actions. For instance, to understand whether a cup is full of coffee, an agent, equipped with a camera, needs to turn on the light and look at the cup from the top. When the proper situations to perceive the desired properties are unknown, an agent needs to learn them and plan to get in such situations. In this paper, we devise a general method to solve this problem by evaluating the confidence of a neural network online and by using symbolic planning. We experimentally evaluate the proposed approach on several synthetic datasets, and show the feasibility of our approach in a real-world scenario that involves noisy perceptions and noisy actions on a real robot. We also note that the contributions of [Lamanna et al., 2023a] and [Lamanna et al., 2023b] are also part of Leonardo Lamanna's

PhD "Integrating Planning and Learning for Agents Acting in Unknown Environments" [Lamanna, 2023], which was the winner of the award for the best AI 2023 doctoral dissertation (Marco Cadoli prize) - awarded by AI*IA - the Italian Association for Artificial Intelligence.

Task 5.5: Synergies Industry, Challenges, Roadmap Concerning on Autonomous Actions in AI Systems

State of the Art. Research in WP5 aims at empowering the agent with the ability of deliberating on how to act in the world in an autonomous fashion without the direct intervention of humans. Crucially, empowering an AI agent with the ability to self-deliberate its own behaviour carries significant risks of the agent getting out-of-control, hence this ability must be balanced with safety. Assessing safety is essential, and formal verification, model checking and automated synthesis to guarantee safety specifications is central to this effort. This line of research involves several fields of AI, including planning, knowledge representation, logics in AI and probabilistic reasoning as well as verification and automated synthesis in Formal Methods. A current shortcoming in the domain is the gap between theoretical research in planning and industry applications.

Achievements. Several interesting research avenues have been identified by WP5 partners during their research, workshops and meetings. In particular, the following areas are considered important:

- Learning action models (related to WP4);
- Non-Markovian reinforcement learning (e.g. reward machines, temporally extended rewards and dynamics);
- Integrating logic-based reasoning about actions and data-driven learning;
- Learning and acting in robotics (behaviour trees);
- Theory of mind in order to reason about beliefs, capabilities and goals, when deliberating and executing actions (related to WP6);
- Connections and synergies with formal methods;
- Goal reasoning and formation;
- Learning and exploiting automata/goal structure;
- Considering multiple models to handle various levels of contingencies.

Clearly, a main research direction concerns the integration and development of model-based and model-free approaches for learning and planning.

These areas were identified in the TAILOR Strategic Research and Innovation Roadmap (SRIR) as impactful areas for European industry. They in particular concern mobility, production, interacting with humans, fintech, entertainment, and many others. For example, autonomous mobile robot platforms are focusing less on hardware aspects and more on organisation and software, to automate warehouses and logistics. This shift is an opportunity for introducing advanced forms of autonomy based on the kind of work done in WP5. Smart manufacturing could benefit from research in learning and reasoning on how to act by automated program-synthesis and learning how to handle unexpected exceptions. Interaction with humans requires autonomous capability in acting in order not to be too annoying to the humans themselves. FinTech is interested in creating autonomous agents that can act rationally while learning from actual data during operation. Also video games, augmented reality, interactive entertainment is heavily relying on these techniques for improving the interaction and the behaviour of avatars.

As also pointed out in the SRIR, learning and reasoning on how to act is strongly connected with other scientific disciplines outside AI. Acting and planning tools have the potential to boost research and technological development e.g. in formal methods, MDPs, best-effort synthesis, operations research, and cybersecurity. There are also significant connections with the humanities.

Task 5.6: Fostering the AI Scientific Community on the Theme of Deciding and Learning How to Act

The theme of deciding and learning how to act has received considerable attention in the scientific community over the past few years. Papers on this topic now appear regularly at top international AI conferences such as IJCAI, AAAI, ECAI, NeurIPS, and ICML. It also features prominently at specialised conferences such as (for example) the *International Conference on Representation Learning (ICLR)*, or the AAAI Fall Symposium 2022 was concerned with this theme in their workshop *Thinking fast and slow and other cognitive theories.*

In the context of TAILOR WP5, the following initiatives deserve particular mention: Both as Program co-Chair (2021) and General Chair (2022) of the International Conference on Principles of Knowledge Representation and Reasoning, Gerhard Lakemeyer has fostered a special track on knowledge representation and machine learning that offers a forum for researchers interested in deciding and learning how to act to present their work at the premier international conference on knowledge representation and reasoning. A number of people from the TAILOR network served as members of the program committee or as area chairs. In particular, Luc De Raedt served as co-Chair of the special track in 2021.

Gerhard Lakemeyer, together with Fredrik Heintz and Sheila McIlraith, organised a Dagstuhl seminar on Cognitive Robotics in September 2022. This seminar brought together leading experts in the field of cognitive robotics, knowledge representation, machine learning, and natural language understanding, among others, and included members from the TAILOR network. The themes of the workshop included Cognitive Robotics and Knowledge Representation, Verification of Robotic Systems, Human-Robot Interaction and Ethics, and Planning and Machine Learning. The latter was led by Hector Geffner, member of TAILOR, and featured spotlight talks as well as group discussions. One of the tangible outcomes of the seminar was a collection of challenge problems and a roadmap for future research. These findings have been published in a Dagstuhl report.

The research groups in TAILOR WP5 organised several events, among them we like to mention the joint workshop WP4 and WP5 at the Third TAILOR Conference in Siena, featuring an invited talk by Hector Geffner on Learning Planning Domains, and an AAAI 2023 Spring Symposium *On the Effectiveness of Temporal Logics on Finite Traces in AI, and many others*, organised by WP5 Giuseppe De Giacomo's group. Moreover TAILOR scientists presented the work done within WP5 in several venues. For example at ECAI2023, WP5 was prominent in several workshops such as *The Verifying Learning AI Systems (VeriLearn) Workshop,* featuring an invited talk by Giuseppe De Giacomo on *"Towards Framed Autonomy"* and the *International Workshop on Logical Aspects in Multi-Agent Systems and Strategic Reasoning 2023,* featuring several papers from WP5 participants. Also the NeurIPS 2023 Workshop on *Generalization in Planning*, though independently organised, was centred on the themes of WP5, and Giuseppe De Giacomo and Hector Geffner from TAILOR WP5 gave two keynotes on *Logic, Automata, and Games in Linear Temporal Logics on Finite Traces*, and on *Learning General Policies and Sketches,* respectively.

Finally we recall that TAILOR developed the concept of the European Summer School on AI (ESSAI) to foster a diverse community of young AI scientists and promote hybridization of knowledge. TAILOR then asked EurAI to handle the ESSAI series to allow the idea to outlast the end of the TAILOR project. We would like to mention that the themes of WP5 were present in several courses of the first edition of ESSAI, ESSAI 2023 and that they are present in several of the submitted courses for ESSAI 2024.

WP5 Publications

An updated list of all publications related to the WP5 is always available online at https://sites.google.com/diag.uniroma1.it/ict-48-tailor-wp5/papers.

[Abate et al., 2021] Abate, A., Gutierrez, J., Hammond, L., Harrenstein, P.,Kwiatkowska, M., Najib, M., Perelli, G., Steeples, T., and Wooldridge, M. (2021). Rational verification: game-theoretic verification of multi-agent systems. Appl. Intell., 51(9):6569–6584.

[Agostinelli et al., 2021] Agostinelli, S., Bergami, G., Fiorenza, A., Maggi, F., Marrella, A., and Patrizi, F. (2021). Discovering declarative process model behaviour from event logs via model learning. In ICPM, pages 48–55. IEEE.

[Agostinelli et al., 2023] Agostinelli, S., Chiariello, F., Maggi, F.M., Marrella, A., and Patrizi, F. (2023). Process Mining Meets Model Learning: Discovering Deterministic Finite State Automata from Event Logs for Business Process Analysis. Inf. Syst. 114: 102180

[Aguilera-Ventura et al., 2023] Aguilera-Ventura, C., Herzig, A., Liu, X., and Lorini, E. (2023) Counterfactual Reasoning via Grounded Distance. In KR.

[Aineto et al., 2021] Aineto, D., Jim´enez, S., and Onaindia, E. (2021). Generalized temporal inference via planning. In KR, pages 22–31.

[Alechina et al., 2022] Alechina, N., De Giacomo, G., Logan, B., and Perelli, G. (2022). Automatic synthesis of dynamic norms for multi-agent systems. In KR.

[Alman et al., 2022] Alman, A., Maggi, F., Montali, M., Patrizi, F., and Rivkin, A. (2022). Multi-model monitoring framework for hybrid process specifications. In CAISE.

[Aminof et al., 2020] Aminof, B., De Giacomo, G., Lomuscio, A., Murano, A., and Rubin, S. (2020). Synthesizing strategies under expected and exceptional environment behaviors.. In IJCAI.

[Aminof et al., 2021a] Aminof, B., De Giacomo, G., Lomuscio, A., Murano, A., and Rubin, S. (2021a). Synthesizing Best-Effort strategies under multiple environment specifications. In KR.

[Aminof et al., 2021b] Aminof, B., De Giacomo, G., and Rubin, S. (2021b). Best-effort synthesis: Doing your best is not harder than giving up. In IJCAI.

[Aminof et al., 2022] Aminof, B., De Giacomo, G., Rubin, S., and Zuleger, F. (2022). Beyond strong-cyclic: Doing your best in stochastic environments. In IJCAI, pages 2525–2531.

[Aminof et al., 2023] Aminof, B., De Giacomo, G., Rubin S., and Zuleger, F. (2023). Stochastic Best-Effort Strategies for Borel Goals. In LICS 2023.

[Aminof et al., 2023a] Aminof, B., De Giacomo, G., Rubin S., (2023a). Reactive Synthesis of Dominant Strategies. In AAAI 2023

[Aminof et al., 2023b] Aminof, B., De Giacomo, G., Di Stasio, A., Francon, H., Rubin, S. and Zhu S. (2023b). LTLf Synthesis Under Environment Specifications for Reachability and Safety Properties. In EUMAS 2023.

[Banihashemi et al., 2023] Banihashemi, B., De Giacomo, G., and Lesperance, Y. (2023). Abstraction of Nondeterministic Situation Calculus Action Theories. In IJCAI

[Belaid et al., 2022] Belaid, M., Belmecheri, N., Gotlieb, A., Lazaar, N., and Spieker, H. (2022). GEQCA: generic qualitative constraint acquisition. In AAAI.

[Biedenkapp et al., 2022] Biedenkapp, A., Speck, D., Sievers, S., Hutter, F., Lindauer, M., and Seipp, J. (2022). Learning domain-independent policies for open list selection. In PRL Workshop.

[Bolander et al. 2023] Belle, V., Bolander, Th., Herzig, A., Nebel, B. (2023). Epistemic planning: Perspectives on the special issue. Artif. Intell. 316: 103842.

[Bonassi et al., 2023a] Bonassi, L., De Giacomo, G., Favorito, M., Fuggitti, F., Gerevini A. E. and Scala, E. Planning for Temporally Extended Goals in Pure-Past Linear Temporal Logic In ICAPS 2023 - Best Student Paper Award.

[Bonassi et al., 2023b] Bonassi, L., De Giacomo, G., Favorito, M., Fuggitti, F., Gerevini A. E. and Scala, E. FOND Planning for Pure-Past Linear Temporal Logic Goals In ECAI 2023.

[Bonet and Geffner, 2021] Bonet, B. and Geffner, H. (2021). General policies, representations, and planning width. In AAAI.

[Bonet et al., 2023] Bonet, B., Drexler, D., and Geffner, H. (2023). General and Reusable Indexical Policies and Sketches. In GenPlan workshop at NeurIPS.

[Bonet et al., 2024] Bonet, B., Drexler, D., and Geffner, H. (2024). On Policy Reuse: An Expressive Language for Representing and Executing General Policies that Call Other Policies. In ICAPS (to appear).

[Boudou et al., 2021] Boudou, J., Herzig, A., and Troquard, N. (2021). Resource separation in dynamic logic of propositional assignments. J. Log. Algebraic Methods Program., 121:100683.

[Bourel et al., 2023] Bourel, H., Jonsson, A., Maillard, O.-A. and Talebi, M. S. (2023). Exploration in Reward Machines with Low Regret. In AISTATS, pages 4114-4146.

[Brunori et al., 2021] Brunori, D., Colonnese, S., Cuomo, F., and Iocchi, L. (2021). A reinforcement learning environment for multi-service UAV-enabled wireless systems.

[Büchner et al., 2021] Büchner, C., Keller, T., and Helmert, M. (2021). Exploiting cyclic dependencies in landmark heuristics. In ICAPS. Runner-Up for Best Student Paper Award at ICAPS 2021.

[Büchner et al., 2023] Büchner, C., Keller, T., Eriksson, S., and Helmert, M. (2023). Landmark Progression in Heuristic Search. In ICAPS. Best Paper Award at ICAPS 2023.

[Calvanese et al., 2022] Calvanese, D., De Giacomo, G., Montali, M., and Patrizi, F. (2022). Verification and monitoring for first-order LTL with persistence-preserving quantification over finite and infinite traces. In IJCAI, pages 2553–2560.

[Chiariello et al., 2022a] Chiariello, F., Maggi, F., and Patrizi, F. (2022a). ASP-based declarative process mining. In AAAI.

[Chiariello et al., 2022b] Chiariello, F., Maggi, F., and Patrizi, F. (2022b). ASP-based declarative process mining (extended abstract). In Proceedings of the 38th International Conference on Logic Programming (Technical Communications) (ICLP). Electronic Proceedings in Theoretical Computer Science (EPTCS).

[Chiariello et al., 2023] Chariello, F., Maggi, F., and Patrizi, F., (2023). From LTL on Process Traces to Finite-State Automata. In BPM (Demos / Resources Forum)

[Christen et al., 2022] Christen, R., Eriksson, S., Pommerening, F., and Helmert, M. (2022). Detecting unsolvability based on separating functions. In ICAPS, pages 44–52.

[Cima et al., 2021] Cima, G., Croce, F., and Lenzerini, M. (2021). Query definability and its approximations in ontology-based data management. In CIKM, pages 271–280. ACM.

[Cipollone et al., 2022] Cipollone, R., De Giacomo, G., Favorito, M., Iocchi, L., and Patrizi, F. (2022). Exploiting multiple levels of abstractions in episodic RL via reward shaping. In PRL Workshop.

[Cipollone et al., 2023] Cipollone, R., Jonsson, A., Ronca, A. and Talebi, M. S. (2023). Provably Efficient Offline Reinforcement Learning in Regular Decision Processes. In NeurIPS.

[Cipollone et al., 2023b] Cipollone, R., De Giacomo, G., Favorito, M., Iocchi, L., and Patrizi, F. (2023b). Exploiting Multiple Abstractions in Episodic RL via Reward Shaping. In AAAI

[Console et al., 2021] Console, M., De Giacomo, G., Lenzerini, M., and Namici, M. (2021). Intensional and extensional views in DL-Lite ontologies. In IJCAI.

[Cooper et al., 2021] Cooper, M. C., Herzig, A., Maffre, F., Maris, F., Perrotin, E., and Régnier, P. (2021). A lightweight epistemic logic and its application to planning. Artificial Intelligence, page 103437.

[Correa et al., 2021] Corrˆea, A. B., Franc`es, G., Pommerening, F., and Helmert, M. (2021). Delete-relaxation heuristics for lifted classical planning. In ICAPS.

[Correa et al., 2022] Corrˆea, A. B., Pommerening, F., Helmert, M., and Franc`es, G. (2022). The FF heuristic for lifted classical planning. In AAAI, pages 9716–9723.

[Correa and Seipp, 2022] Corrˆea, A. B. and Seipp, J. (2022). Best-first width search for lifted classical planning. In ICAPS, pages 11–15.

[Correa et al., 2023] Corrêa, A. B., Hecher, M., Helmert, M., Longo, D. M., Pommerening, F., and Woltran, S. (2023). Grounding Planning Tasks Using Tree Decompositions and Iterated Solving. In ICAPS. Runner-Up for Best Student Paper Award at ICAPS 2023.

[Correa et al., 2024] Corrˆea, A., B., De Giacomo, G., Helmert, M., and Rubin, S. (2024). Planning with Object Creation.. ICAPS 2024. To appear.

[de A. Santos et al., 2021] de A. Santos, L. R., Meneguzzi, F., Fraga Pereira, R., and Pereira, A. G. (2021). An lp-based approach for goal recognition as planning. In AAAI.

[De Giacomo et al., 2019] De Giacomo, G., Favorito, M., Iocchi, L., and Patrizi, F. (2019). Foundations for Restraining Bolts: Reinforcement Learning with LTLf/LDLf Restraining Specifications. In ICAPS

[De Giacomo et al., 2020a] De Giacomo, G., Di Stasio, A., Fuggitt, F., and Rubin S. IJCAI Pure-Past Linear Temporal and Dynamic Logic on Finite Traces 2020 (Survey Track).

[De Giacomo et al., 2020b] De Giacomo, G., Favorito, M., Iocchi, L., and Patrizi, F. (2020b). Imitation Learning over Heterogeneous Agents with Restraining Bolts. In ICAPS

[De Giacomo et al., 2021a] De Giacomo, G., Di Stasio, A., Perelli, G., and Zhu, S. (2021a). Synthesis with mandatory stop actions. In KR.

[De Giacomo et al., 2021b] De Giacomo, G., Di Stasio, A., Tabajara, L. M., Vardi, M., and Zhu, S. (2021b). Finite-trace and generalized-reactivity specifications in temporal synthesis. In IJCAI.

[De Giacomo and Favorito, 2021] De Giacomo, G. and Favorito, M. (2021). Compositional approach to translate ltlf/ldlf into deterministic finite automata. In ICAPS, volume 14.

[De Giacomo et al., 2022a] De Giacomo, G., Favorito, M., and Fuggitti, F. (2022a). Planning for temporally extended goals in pure-past linear temporal logic: A polynomial reduction to standard planning. CoRR, abs/2204.09960.

[De Giacomo et al., 2021c] De Giacomo, G., Favorito, M., Iocchi, L., and Patrizi, F. (2021c). Domain-Independent reward machines for modular integration of planning and learning. In PRL Workshop.

[De Giacomo et al., 2021f] De Giacomo, G., Murano, A., Patrizi, F., and Perelli, G. (2021f). Timed Trace Alignment with Metric Temporal Logic over Finite Traces. In KR

[De Giacomo et al., 2022b] De Giacomo, G., Favorito, M., Leotta, F., Mecella, M., and Silo, L. (2022b). Modeling resilient cyber-physical processes and their composition from digital twins via markov decision processes. CEUR.

[De Giacomo et al., 2022c] De Giacomo, G., Favorito, M., Li, J., Vardi, M. Y., Xiao, S., and Zhu, S. (2022c). Ltlf synthesis as AND-OR graph search: Knowledge compilation at work. In IJCAI, pages 2591–2598.

[De Giacomo et al., 2022d] De Giacomo, G., Favorito, M., Mecella, M., Leotta, F., and Silo, L. (2022d). Digital twins composition via markov decision processes. In ITAL-IA Workshop.

[De Giacomo et al., 2022e] De Giacomo, G., Felli, P., Logan, B. and Patrizi, F., (2022e) Controller Synthesis for Manufacturing Systems in the Situation Calculus. In IJCAI – journal track.

[De Giacomo et al., 2022f] De Giacomo, G., De Masellis, R., Maggi, F. M., and Montali, M. (2022f). Monitoring Constraints and Metaconstraints with Temporal Logics on Finite Traces. ACM Transactions on Software Engineering and Methodology, 31(4):68, 2022.

[De Giacomo et al., 2021d] De Giacomo, G., Felli, P., Logan, B., Patrizi, F., and Sardi˜na, S. (2021d). Situation calculus for controller synthesis in manufacturing systems with first-order state representation. Artificial Intelligence.

[De Giacomo et al., 2021e] De Giacomo, G., Felli, P., Montali, M., and Perelli, G. (2021e). HyperLDLf: a logic for checking properties of finite traces process logs. In IJCAI.

[De Giacomo and Lespérance, 2021] De Giacomo, G. and Lesp ́erance, Y. (2021). The nondeterministic situation calculus. In KR.

[De Giacomo et al., 2021g] De Giacomo, G., Oriol, X., Rosati, R., and Savo, D. F. (2021g). Instance-level update in dl-lite ontologies through first-order rewriting. JAIR, 70:1335–1371.

[De Giacomo and Perelli, 2023] De Giacomo, G. and Perelli, G. (2023). Behavioral QLTL. In EUMAS.

[De Giacomo et al., 2023] De Giacomo, G., Di Stasio, G., Tabajara L. M., Vardi M. Y., and Zhu, S. (2023). Finite-trace and Generalized-Reactivity Specifications in Temporal Synthesis. In Formal Methods in System Design. 2023.

[De Giacomo et al., 2023a] De Giacomo, G., Parretti, G., and Zhu, S. (2023a). LTLf Best-Effort Synthesis in Nondeterministic Planning Domains. In ECAI

[De Giacomo et al., 2023b] De Giacomo G., Parretti, G., and Zhu, S. (2023b). Symbolic LTLf Best-Effort Synthesis. In EUMAS

[De Giacomo et al., 2023c] De Giacomo, G., Fuggitt, F., Maggi F. M., Marrella A., Favorito, M. and Patrizi, F. (2023c). A Tool for Declarative Trace Alignment via Automated Planning. In Software Impacts 16, 2023.

[De Giacomo et al., 2023d] De Giacomo, G., Favorito, M., Leotta, F., Mecella, M., Monti, F., and Silo, L. (2023d). AIDA: A Tool for Resiliency in Smart Manufacturing. CAiSE Forum 2023.

[De Giacomo et al., 2023e] De Giacomo, G., Favorito, M., Leotta, F., Mecella, M., and Silo, L. (2023e). Digital Twin Composition in Smart Manufacturing via Markov Decision Processes. Computers in Industry 149, 2023.

[De Giacomo et al., 2023f] De Giacomo, F., Fried, D., Patrizi, F., and Zhu, S. (2023f). Mimicking Behaviors in Separated Domains. J. Artif. Intell. Res. (JAIR) 77: 1087-1112 (2023)

[Drexler et al., 2021] Drexler D., Seipp, J. and Speck, D. (2021). Subset-saturated transition cost partitioning. In ICAPS, pages 131–139.

[Drexler et al., 2021] Drexler, D., Seipp, J., and Geffner, H. (2021). Expressing and exploiting the common subgoal structure of classical planning domains using sketches. In KR.

[Drexler et al., 2022] Drexler, D., Seipp, J., and Geffner, H. (2022). Learning sketches for decomposing planning problems into subproblems of bounded width. In ICAPS.

[Drexler et al., 2023] Drexler, D., Seipp, J., and Geffner, H. (2023). Learning Hierarchical Policies by Iteratively Reducing the Width of Sketch Rules. In KR.

[Drexler et al., 2024] Drexler, D., Seipp, J., and Geffner, H. (2024). Expressing and Exploiting Subgoal Structure in Classical Planning Using Sketches. In ICAPS (to appear).

[Engesser et al. 2024] Engesser, Th., Herzig, A., Perrotin, E. (2024). Towards Epistemic-Doxastic Planning with Observation and Revision. In AAAI (to appear)

[Favorito and Zhu, 2023] Favorito, M. and Zhu, S. (2023). LydiaSyft: A Symbolic Compositional Synthesizer for LTLf Specifications. In SYNTHCOMP, 2023.

[Favorito, 2023] Favorito, M. (2023). Forward LTLf Synthesis: DPLL At Work. In IPS-RCRA-SPIRIT@AI*IA 2023

[Feng et al. 2023] Feng, Q. (2023). A Logic of Only-Believing over Arbitrary Probability Distributions. AAMAS 2023: 355-363.

[Ferber et al., 2022a] Ferber, P., Geißer, F., Trevizan, F., Helmert, M., and Hoffmann, J. (2022). Neural network heuristic functions for classical planning: Bootstrapping and comparison to other methods. In ICAPS, pages 583–587.

[Ferber et al., 2022b] Patrick Ferber, Liat Cohen, J. S. and Keller, T. (2022). Learning and exploiting progress states in greedy best-first search. In IJCAI, pages 4740–4746.

[Ferber and Seipp, 2022] Ferber, P. and Seipp, J. (2022). Explainable planner selection for classical planning. In AAAI, pages 9741–9749. AAAI Press.

[Fraga Pereira et al., 2021] Fraga Pereira, R., Fuggitti, F., and De Giacomo, G. (2021). Recognizing LTLf/PLTLf goals in fully observable non-deterministic domain models. CoRR, abs/2103.11692.

[Fraga Pereira et al., 2022] Fraga Pereira, R., Messa, F., Grahl Pereira, A., and De Giacomo, G. (2022). Iterative depth-first search for fully observable non-deterministic planning. In ICAPS.

[Fraga Pereira et al., 2024] Pereira, R. F., Fuggitti, F., Meneguzzi, F., & De Giacomo, G. (2024). Temporally extended goal recognition in fully observable non-deterministic domain models. Appl. Intell. 54(1): 470-489 (2024).

[Francès et al., 2021] Franc`es, G., Bonet, B., and Geffner, H. (2021). Learning general planning policies from small examples without supervision. In AAAI, pages 11801–11808.

[Furelos-Blanco et al., 2023] Furelos-Blanco, D., Law, M., Jonsson, A., Broda, K. and Russo, A. (2023). Hierarchies of reward machines. In ICML, pages 10494-10541.

[Geffner, 2022] Target Languages (vs. Inductive Biases) for Learning to Act and Plan. Geffner, H. In AAAI, pages 12326–12333, 2022. AAAI Press

[Geisler et al., 2022] Geisler, S., Vidal, M., Cappiello, C., L´oscio, B. F., Gal, A., Jarke, M., Lenzerini, M., Missier, P., Otto, B., Paja, E., Pernici, B., and Rehof, J. (2022).

Knowledge-driven data ecosystems toward data transparency. ACM J. Data Inf. Qual., 14(1):3:1–3:12.

[Gutierrez et al., 2021a] Gutierrez, J., Harrenstein, P., Perelli, G., and Wooldridge, M. (2021a). Expressiveness and nash equilibrium in iterated boolean games. ACM Trans. Comput. Logic, 22(2).

[Gutierrez et al., 2021b] Gutierrez, J., Murano, A., Perelli, G., Rubin, S., Steeples, T., and Wooldridge, M. (2021b). Equilibria for games with combined qualitative and quantitative objectives. Acta Informatica, 58(6):585–610.

[Habering et al., 2021] Habering, D., Hofmann, T., and Lakemeyer, G. (2021). Using platform models for a guided explanatory diagnosis generation for mobile robots. In IJCAI, pages 1908–1914. ijcai.org.

[Heller et al., 2022] Heller, D., Ferber, P., Bitterwolf, J., Hein, M., and Hoffmann, J. (2022). Neural network heuristic functions: Taking confidence into account. In SoCS, pages 223–228.

[Helmert et al., 2022] Helmert, M., Sievers, S., Rovner, A., and Corrˆea, A. B. (2022). On the complexity of heuristic synthesis for satisficing classical planning: Potential heuristics and beyond. In ICAPS, pages 124–133.

[Herzig et al., 2022] Herzig, A., Lorini, E., and Perrotin, E. (2022). A computationally grounded logic of 'seeing-to-it-that'. In Raedt, L. D., editor, Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI 2022, Vienna, Austria, 23-29 July 2022, pages 2648–2654.

[Herzig et al., 2021] Herzig, A., Maris, F., and Perrotin, E. (2021). A dynamic epistemic logic with finite iteration and parallel composition. In KR, pages 676–680.

[Herzig and Yuste-Ginel, 2021a] Herzig, A. and Yuste-Ginel, A. (2021a). Abstract argumentation with qualitative uncertainty: An analysis in dynamic logic. In Logic and Argumentation - 4th International Conference, CLAR 2021, Hangzhou, China, October 20-22, 2021, Proceedings, volume 13040 of Lecture Notes in Computer Science, pages 190–208. Springer.

[Herzig and Yuste-Ginel, 2021b] Herzig, A. and Yuste-Ginel, A. (2021b). On the epistemic logic of incomplete argumentation frameworks. In KR, pages 681–685.

[Herzig and Yuste-Ginel, 2021c] Herzig, A. and Yuste-Ginel, A. (2021c). Multi-Agent Abstract Argumentation Frameworks With Incomplete Knowledge of Attacks. In IJCAI, pages 681–685.

[Hu and Speck, 2022] Hu, K. and Speck, D. (2022). On bidirectional heuristic search in classical planning: An analysis of bae*. In SoCS, pages 91–99.

[Infante et al., 2023] Infante, G., Jonsson, A. and Gómez, V. (2023). Optimal Hierarchical Average-Reward Linearly-solvable Markov Decision Processes. In EWRL.

[Ielo et al., 2023] Ielo, A., Law, M., Fionda, V., Ricca, F., De Giacomo, G., and Russo, A. (2023).Towards ILP-Based LTLf Passive Learning. ILP 2023 - Best Paper Award.

[Kin Max Gusm˜ao, 2021] Kin Max Gusm˜ao, Ramon Fraga Pereira, F. M. (2021). Inferring agents preferences as priors for probabilistic goal recognition. In SPARK Workshop (ICAPS).

[Kloßner et al., 2022] Klößner, T., Pommerening, F., Keller, T., and Röger, G. (2023). Cost Partitioning Heuristics for Stochastic Shortest Path Problems. In ICAPS. Runner-Up for Best Student Paper Award at ICAPS 2022.

[Kuric et al., 2024] Kuric, D., Infante, G., Gómez, V., Jonsson, A. and van Hoof, H. (2024). Planning with a Learned Policy Basis to Optimally Solve Complex Tasks. In ICAPS.

[Lamanna et al., 2021a] Lamanna, L., Saetti, A., Serafini, L., Gerevini, A. E., and Traverso, P. (2021a). Online learning of action models for PDDL planning. In IJCAI.

[Lamanna et al., 2021b] Lamanna, L., Serafini, L., Gerevini, A. E., Saetti, A., and Traverso, P. (2021b). On-line learning of planning domains from sensor data in PAL: Scaling up to large state spaces. In AAAI.

[Lamanna et al., 2022] Lamanna, L., Serafini, L., Saetti, A., Gerevini, A., and Traverso, P. (2022). Online grounding of symbolic planning domains in unknown environments. In KR.

[Lamanna et al. 2023a] Leonardo Lamanna, Luciano Serafini, Mohamadreza Faridghasemnia, Alessandro Saffiotti, Alessandro Saetti, Alfonso Gerevini, Paolo Traverso: Planning for Learning Object Properties. AAAI 2023: 12005-12013

[Lamanna et al. 2023b] Leonardo Lamanna, Mohamadreza Faridghasemnia, Alfonso Gerevini, Alessandro Saetti, Alessandro Saffiotti, Luciano Serafini, Paolo Traverso: Learning to Act for Perceiving in Partially Unknown Environments. IJCAI 2023: 5485-5493

[Lamanna et al., 2023] Leonardo Lamanna. Integrating Planning and Learning for Agents Acting in Unknown Environments [Doctoral thesis, Università degli Studi di Brescia]

[Lembo et al. 2022] Lembo, D., Santarelli, V., Savo, D.F. and De Giacomo, G., (2022). Graphol: A Graphical Language for Ontology Modeling Equivalent to OWL 2. In Future Internet, 14(3), p.78.

[Lenzerini et al., 2021] Lenzerini, M., Lepore, L., and Poggi, A. (2021). Meta-modeling and metaquerying in OWL 2 QL. Artificial Intelligence, 292:103432.

[Lesperance et al., 2024] Lesperance, Y., De Giacomo, G., Rostamigiv, M., and Khan, S. M. (2024). Abstraction of Situation Calculus Concurrent Game Structures. In AAAI

[Liu and Lakemeyer, 2021] Liu, D. and Lakemeyer, G. (2021). Reasoning about beliefs and meta-beliefs by regression in an expressive probabilistic action logic. In IJCAI, pages 1951–1958. ijcai.org.

[Liu et al., 2023a] Liu, D., Huang, Q., Belle, V., and Lakemeyer, G. (2023). Verifying Belief-Based Programs via Symbolic Dynamic Programming. ECAI 2023: 1497-1504.

[Liu et al., 2023b] Liu, D., Feng, Q., Belle, V., and Lakemeyer, G. (2023). Concerning Measures in a First-order Logic with Actions and Meta-beliefs. KR 2023: 451-460.

[Max Gusmao et al., 2021] Max Gusmao, K., Fraga Pereira, R., Grahl Pereira, A., and Meneguzzi, F. (2021). Inferring agents preferences as priors for probabilistic goal recognition. CoRR, abs/2102.11791.

[Meneguzzi and Fraga Pereira, 2021] Meneguzzi, F. and Fraga Pereira, R. (2021). A survey on goal recognition as planning. In IJCAI.

[Menguy et al., 2022] Menguy, G., Bardin, S., Lazaar, N., and Gotlieb, A. (2022). Automated program analysis: Revisiting precondition inference through constraint acquisition. In IJCAI.

[Micheli and Valentini, 2021] Micheli, A. and Valentini, A. (2021). Synthesis of search heuristics for temporal planning via reinforcement learning. In AAAI.

[Mittelmann et al., 2021] Mittelmann, M., Herzig, A., and Perrussel, L. (2021). Epistemic reasoning about rationality and bids in auctions. In JELIA, volume 12678 of Lecture Notes in Computer Science, pages 116–130. Springer.

[Mittelmann et al., 2022a] Mittelmann, M., Maubert, B., Murano, A., and Perrussel, L. (2022a). Automated synthesis of mechanisms. In IJCAI, pages 426–432. ijcai.org.

[Mittelmann et al., 2022b] Mittelmann, M., Maubert, B., Murano, A., and Perrussel, L. (2022b). Synthesis of mechanisms with strategy logic (short paper). In ICTCS, CEUR Workshop Proceedings. CEUR-WS.org.

[Oriol et al., 2021] Oriol, X., De Giacomo, G., Esta˜nol, M., and Teniente, E. (2021). Embedding reactive behavior into artifact-centric business process models. FGCS, 117:97–110.

[Parr and Pezzulo, 2021] Parr, T. and Pezzulo, G. (2021). Understanding, explanation, and active inference. Frontiers in Systems Neuroscience, 15.

[Pommerening et al., 2021] Pommerening, F., Keller, T., Halasi, V., Seipp, J., Sievers, S., and Helmert, M. (2021). Dantzig-wolfe decomposition for cost partitioning. In ICAPS.

[Ramon Fraga Pereira, 2021] Ramon Fraga Pereira, Nir Oren, F. M. (2021). Landmark-based approaches for goal recognition as planning. In ICAPS, Journal Track.

[Rodriguez et al., 2021a] Rodriguez, I. D., Bonet, B., Romero, J., and Geffner, H. (2021a). Learning first-order representations for planning from black-box states: New results. In KR.

[Rodriguez et al., 2021b] Rodriguez, I. D., Bonet, B., Sardi˜na, S., and Geffner, H. (2021b). Flexible FOND planning with explicit fairness assumptions. In ICAPS, volume 31, pages 290–298.

[Ronca and De Giacomo, 2021a] Ronca, A. and De Giacomo, G. (2021a). Efficient PAC reinforcement learning in regular decision processes. In PRL Workshop.

[Ronca and De Giacomo, 2021b] Ronca, A. and De Giacomo, G. (2021b). Efficient PAC reinforcement learning in regular decision processes. In IJCAI.

[Ronca et al., 2022a] Ronca, A., Kaminski, M., Grau, B. C., and Horrocks, I. (2022a). The delay and window size problems in rule-based stream reasoning. Artificial Intelligence, page 103668.

[Ronca et al., 2022b] Ronca, A., Paludo Licks, G., and De Giacomo, G. (2022b). Markov abstractions for PAC reinforcement learning in non-Markov decision processes. In IJCAI.

[Ronca et al., 2023] Ronca, A., Knorozova, N. A., and De Giacomo, G. (2023). Automata Cascades: Expressivity and Sample Complexity. AAAI 2023

[Rosa Amado et al., 2021] Rosa Amado, L., Fraga Pereira, R., and Meneguzzi, F. (2021). Combining LSTMs and symbolic approaches for robust plan recognition. In AAMAS.

[Segovia-Aguas et al., 2021] Segovia-Aguas, J., Jiménez, S., and Jonsson, A. (2021). Generalized planning as heuristic search. In ICAPS, volume 31, pages 569–577.

[Segovia-Aguas et al., 2022a] Segovia-Aguas, J., Jiménez Celorrio, S., Sebastiá, L., and Jonsson, A. (2022a). Scaling-up generalized planning as heuristic search with landmarks. In SoCS, pages 171–179.

[Segovia-Aguas et al., 2022b] Segovia-Aguas, J., Jiménez Celorrio, S., and Jonsson, A. (2022b). Computing programs for generalized planning as heuristic search (extended abstract). In IJCAI-22, pages 5334–5338. Sister Conferences Best Papers.

[Segovia-Aguas et al., 2024] Segovia-Aguas, J., Jiménez Celorrio, S. and Jonsson, A. (2024). Generalized Planning as Heuristic Search: A new planning search-space that leverages pointers over objects. Artificial Intelligence, volume 330.

[Seipp, 2021] Seipp, J. (2021). Online saturated cost partitioning for classical planning. In ICAPS.

[Seipp et al., 2021] Seipp, J., Keller, T., and Helmert, M. (2021). Saturated post-hoc optimization for classical planning. In AAAI.

[Sievers et al., 2022] Sievers, S., Gnad, D., and Torralba, A. (2022). Additive pattern databases for decoupled search. In SOCS. Best Paper Award at SoCS 2022.

[Sievers and Helmert, 2021] Sievers, S. and Helmert, M. (2021). Merge-and-shrink: A compositional theory of transformations of factored transition systems. JAIR, 71:781–883.

[Sievers and Wehrle, 2021] Sievers, S. and Wehrle, M. (2021). On weak stubborn sets in classical planning. In IJCAI, pages 4167–4174.

[Singh et al., 2021] Singh, A., Lipovetzky, N., Ramirez, M., and Segovia-Aguas, J. (2021). Approximate novelty search. In ICAPS, volume 31, pages 349–357.

[Speck and Seipp, 2022] Speck, D. and Seipp, J. (2022). New refinement strategies for cartesian abstractions. In ICAPS, pages 348–352.

[Ståhlberg et al., 2022a] Ståhlberg, S., Bonet, B., and Geffner, H. (2022a). Learning general optimal policies with graph neural networks: Expressive power, transparency, and limits. In ICAPS, pages 629–637.

[Ståhlberg et al., 2022b] Ståhlberg, S., Bonet, B., and Geffner, H. (2022b). Learning generalized policies without supervision using GNNs. In KR.

[Ståhlberg et al., 2021] Ståhlberg, S., Francès, G., and Seipp, J. (2021). Learning generalized unsolvability heuristics for classical planning. In IJCAI, pages 4175–4181.

[Ståhlberg et al., 2023] Ståhlberg, S., Bonet, B., and Geffner, H. (2023). Learning General Policies with Policy Gradient Methods. In KR.

[Steccanella et al., 2023] Steccanella, L., Totaro, S. and Jonsson, A. (2023). Hierarchical representation learning for Markov decision processes. In CoLLAs, pages 568-585.

[Steinmetz et al., 2022] Steinmetz, M., Fiˇser, D., Eniser, H. F., Ferber, P., Gros, T., Heim, P., Höller, D., Schuler, X., Wüstholz, V., Christakis, M., and Hoffmann, J. (2022). Debugging a policy: Automatic action-policy testing in AI planning. In ICAPS, pages 353–361.

[Suárez-Hernández et al., 2021a] Suárez-Hernández, A., Andriella, A., Taranović, A., Segovia-Aguas, J., Torras, C., and Alenyà, G. (2021a). Automatic learning of cognitive exercises for socially assistive robotics. In 2021 30th IEEE International

Conference on Robot & Human Interactive Communication (RO-MAN), pages 139–146. IEEE.

[Suárez-Hernández et al., 2021b] Suárez-Hernández, A., Segovia-Aguas, J., Torras, C., and Alenyà, G. (2021b). Online action recognition. In AAAI.

[Tiger et al., 2021] Tiger, M., Bergstrom, D., Norrstig, A., and Heintz, F. (2021). Enhancing lattice-based motion planning with introspective learning and reasoning. IEEE Robotics and Automation Letters, 6(3):4385–4392.

[Torralba et al., 2021] Torralba, A., Seipp, J., and Sievers, S. (2021). Automatic instance generation for classical planning. In ICAPS.

[Trapasso et al., 2023] Trapasso, A., Santilli, S., Iocchi, L., & Patrizi, F. (2023). A formalization of multi-agent planning with explicit agent representation. In ACM,

[Umili et al., 2021] Umili, E., Antonioni, E., Riccio, F., Capobianco, R., Nardi, D., and De Giacomo, G. (2021). Learning a symbolic planning domain through the interaction with continuous environments. In PRL Workshop.

[Varricchione et al. 2024] Varricchione, G., Alechina, N., Dastani, M., De Giacomo, G., Logan, B., and Perelli, G. (2024). Pure-Past Action Masking. In AAAI

[Viehmann et al., 2021] Viehmann, T., Hofmann, T., and Lakemeyer, G. (2021). Transforming robotic plans with timed automata to solve temporal platform constraints. In IJCAI, pages 2083–2089.

[von Tschammer et al., 2022] von Tschammer, J., Mattm ̈uller, R., and Speck, D. (2022). Loopless top-k planning. In ICAPS, pages 380–384.

[Umili et al., 2023] Umili, E., Capobianco, R., and De Giacomo, G. (2023). Grounding LTLf Specifications in Image Sequences. In KR 2023

[Xiao et al., 2021] Xiao, S., Li, J., Zhu, S., Shi, Y., Pu, G., and Vardi, M. (2021). On-the-fly synthesis for LTL over finite traces. In AAAI.

[Yang et al., 2021] Yang, W.-C., Raskin, J.-F., and De Raedt, L. (2021). Lifted model checking for relational MDPs. Machine Learning.

[Yuste-Ginel, A. and Herzig, A. (2023). Qualitative uncertainty and dynamics of argumentation through dynamic logic. J. Log. Comput. 3(2).

[Zhu and De Giacomo, 2022a] Zhu, S. and De Giacomo, G. (2022a). Act for your duties but maintain your rights. In KR.

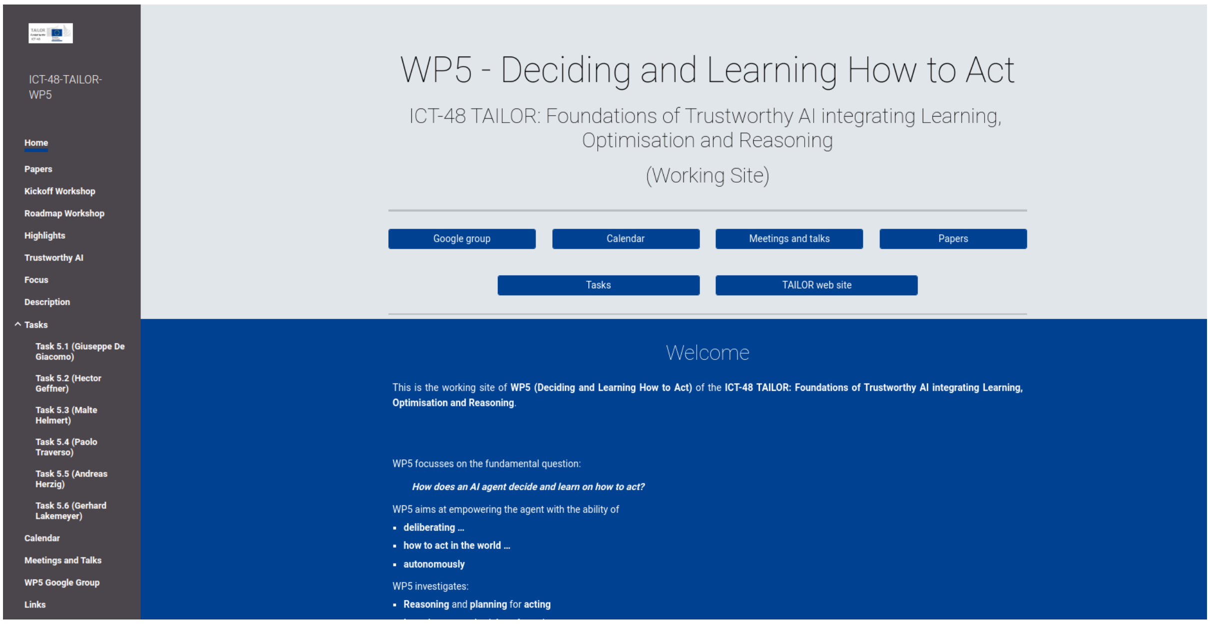
[Zhu and De Giacomo, 2022b] Zhu, S. and De Giacomo, G. (2022b). Synthesis of maximally permissive strategies for ltlf specifications. In IJCAI, pages 2783–2789.

[Zhu et al., 2021] Zhu, S., Tabajara, L. M., Pu, G., and Vardi, M. Y. (2021). On the power of automata minimization in temporal synthesis. In GandALF, volume 346 of EPTCS, pages 117–134.

Appendix 1: Online Website

Due to the unprecedented circumstances brought by the pandemic over the last two years, we decided to coordinate every WP5 activity, meeting, and achievement through an online shared platform identified by a working website available at https://sites.google.com/diag.uniroma1.it/ict-48-tailor-wp5.

Screenshot of the main page.



This working website also features a dedicated forum (https://groups.google.com/g/tailor-wp5-open) for people to exchange ideas, information and collaborate on projects related to WP5.

Appendix 2: Workshops, Open Talks and Courses

Workshops

WP5 Kickoff Workshop (10/12/2020)

Program.
14:45 - 15:00 - Gathering
15:00 - 16:00 - Invited talk by Murray Shanahan (Imperial College London, Deep Mind) - chaired by Hector Geffner - video
16:00 - 16:15 - Break
16:15 - 17:45 - Scientific panels for every task chaired by Kristian Kersting - video

- Task 5.1: Extended and Multi-facet Models of the World Dynamics and Tasks – Giuseppe De Giacomo
- Task 5.2: Integrating Data-based Methods with Model-based Methods in Deciding and Learning How to Act – Hector Geffner
- Task 5.3: Learning for Reasoners and Planners, and Reasoners and Planners for Learning – Malte Helmert
- Task 5.4: Monitoring and Controlling to Make Actions AI Trustworthy in the Real World – Paolo Traverso
- Task 5.5: Synergies Industry, Challenges, Roadmap Concerning on Autonomous Actions in AI Systems – Andreas Herzig
- Task 5.6: Fostering the AI Scientific Community on the Theme of Deciding and Learning How to Act – Gerhard Lakemeyer

17:45 - 18:00 - Break
18:00 - 19:00 - Open discussion on how to organise WP activities (Workshops, micro-projects, site, discussion groups, phd/postdocs managed activities, etc.)

WP5 Roadmap Workshop (20/07/2021)

WP5 held a workshop, chaired by Andreas Herzig (leader of Task 5.5), on the WP5 contribution to the TAILOR roadmap (WP2).

Program.
14:00 - 15:00 - Invited talk by Sheila McIlraith (University of Toronto, ON, Canada)
15:00 - 15:30 - Presentation of the roadmap by Marc Schoenauer
15:30 - 16:00 - Breakout session
16:00 - 17:00 - Presentation of the breakout session groups and general discussion

Joint WP4-WP5 workshop (04/06/2023)

WP4 and WP5 organized a join workshop charred by Luc de Readt and Giuseppe De Giacomo within the 3rd TAILOR conference held on5-6 June 2023 in Siena, Italy.

Program.
16:00-17:00 Invited talk by Hector Geffner
17:00-17:30 Quick presentations of the work in the two WPs
17:30-19:00 Brainstorming among WP4 and WP5 participants

Open Talks

WP5 has participated to two TAILOR Open Monthly Meetings, presenting two open scientific talks:

- *Deciding and Learning How to Act* by Giuseppe De Giacomo on March 22, 2022
- *Top-down representation learning for acting and planning* by Hector Geffner on September 6, 2022.

Courses

Artificial Intelligence and Social Intelligence (17/07/2022)

Host Institutions: TAILOR Summer School in Barcelona
Instructor: Andreas Herzig (Centre National de Recherche Scientifique (CNRS) Institut de Recherche en Informatique de Toulouse (IRIT) Univ. Toulouse, France)
Link:
https://www.irit.fr/~Andreas.Herzig/Cours/CSocIntell_Epi/202206_Barca4Tailor_Latex/SocintellEpilogic.pdf

Game-Theoretic Approach to Planning and Synthesis (4/7/2022 – 8/7/2022)

Host Institutions: Sapienza University & ICT-48 TAILOR
Instructors: Giuseppe De Giacomo, Antonio Di Stasio, Giuseppe Perelli, Shufang Zhu (Sapienza University of Rome)
Link: https://whitemech.github.io/courses
Description: This course introduces AI planning and program synthesis for tasks (goals) expressed over finite traces instead of states. Specifically, borrowing from Formal Methods, we will consider tasks and environment specifications expressed in LTL and its finite trace variant LTLf. We will review the main results and algorithmic techniques to handle planning in nondeterministic domains. Then, we will draw connections with verification, and reactive

synthesis, together with their game-theoretic solution techniques. The main catch is that working with these logics can be based on devising suitable 2-players games and finding strategies, i.e., plans, to win them. Specifically, we will cover the following topics: Planning in nondeterministic domain, Temporal Logics, LTL, LTLf, Game-theoretic Techniques, Safety Games, Reachability Games, Games for LTL/LTLf objectives, and Reactive Synthesis. This course is partially based on the work carried out in ERC Advanced Grant WhiteMech and EU ICT-48 TAILOR.

This course has become an 20-hour course for the Italian National PhD Program in AI, taught by Giuseppe De Giacomo held in May 2023.
(https://sites.google.com/diag.uniroma1.it/game-based-planning-synthesis/)

Non-Markov Decision Processes and Reinforcement Learning (7/11/2022 – 21/11/2022)

Host Institutions: Sapienza University with the support of ICT-48 TAILOR and AIDA
Instructors: Giuseppe De Giacomo, Luca Iocchi, Fabio Patrizi, Alessandro Ronca (Sapienza University of Rome)
Guest Lecturers: Roberto Cipollone, Gabriel Paludo Licks, Elena Umili (Sapienza University of Rome)
Link: https://whitemech.github.io/courses
Description: This course is on non-Markov decision processes, where rewards and dynamics can depend on the history of events. This is contrast with Markov Decision Processes, where the dependency is limited to the last state and action. We study how to specify non-Markov reward functions and dynamics functions using Linear Temporal Logic on finite traces. The resulting decision processes are called Regular Decision Processes, and we show how to solve them by extending solution techniques for Markov Decision Processes. Then, we turn to Reinforcement Learning. First, we study the Restraining Bolt, a device that enables an agent to learn a specified non-Markov behaviour while relying on the Markov property. Second, we study how an agent can achieve an optimal behaviour in a non-Markov domain, by learning a finite-state automaton that describes rewards and dynamics. Specifically we will cover the following topics: MDP with Non-Markov Rewards, Non-Markov Dynamics, Regular Decision Processes, Restraining Bolts, Linear Time Logic on finite traces as a reward/dynamics specification language, Reinforcement Learning, Deep Reinforcement Learning, Automata Learning. This course is partially based on the work carried out in ERC Advanced Grant WhiteMech and EU ICT-48 TAILOR.

Game-Theoretic Approach to Planning and Synthesis (24/7/2023 – 28/7/2023)

Host intitution: ESSAI 2023 Ljubljana, Slovenia & ICT-48 TAILOR
Instructors: Antonio Di Stasio (University of Oxford), Giuseppe Perelli (Sapienza University of Rome), Shufang Zhu (University of Oxford).
Link: https://essai.si/game-theoretic-approach-to-planning-and-synthesis/
Description: This course introduces AI planning and program synthesis for tasks (goals) expressed over finite traces instead of states. Specifically, borrowing from Formal Methods, we will consider tasks and environment specifications expressed in LTL and its finite trace variant LTLf. We will review the main results and algorithmic techniques to handle planning in nondeterministic domains. Then, we will draw connections with verification, and reactive synthesis, together with their game-theoretic solution techniques. The main catch is that working with these logics can be based on devising suitable 2-players games and finding strategies, i.e., plans, to win them. Specifically, we will cover the following topics: Planning in nondeterministic domain, Temporal Logics, LTL, LTLf, Game-theoretic Techniques, Safety Games, Reachability Games, Games for LTL/LTLf objectives, and Reactive Synthesis. This course is partially based on the work carried out in ERC Advanced Grant WhiteMech and EU ICT-48 TAILOR.

Learning to Act and Plan (24/7/2023 – 28/7/2023)

Host intitution: ESSAI 2023 Ljubljana, Slovenia & ICT-48 TAILOR
Instructors: Blai Bonet (UPF, Barcelona), Hector Geffner (RWTH Aachen University)
Link: https://essai.si/learning-to-act-and-plan/
Description: This is an advanced course on learning to act and plan in different settings: when the action model is known, when it is not known but has to be learned, and when it is not known and doesn't have to be learned at all. The five lectures will be as follows: 1. Intro: models and solvers, model-based solvers vs. model-free learners. Deep learning and stochastic gradient descent as another class of model and solver. 2. Classical planning: languages and algorithms; planning as heuristic search and as SAT, learning planning models. 3. MDPs and RL: the model, basic model-based algorithms; reinforcement learning: model-based and model-free. Policy gradient and policy optimization. 4. General plans: learning policies that generalize across domains. Representing and learning such plans using combinatorial and deep learning approaches. 5. Hierarchies and problem decomposition: width and width-based search; representing problem decomposition in a general language; learning subgoal structure.

Appendix 3: Awards

Some work carried out within the WP5 has been recognized with prestigious awards at top-tier international conferences. In particular, the following papers have been awarded:

- "Learning Generalized Unsolvability Heuristics for Classical Planning". Ståhlberg, S.; Francès, G.; and Seipp, J.. **Best Paper Award at IJCAI 2021**
- "Flexible FOND Planning with Explicit Fairness Assumptions". Rodriguez, I. D; Bonet, B.; Sardiña, S.; and Geffner, H. **Best Paper Award at ICAPS 2021**
- "Exploiting Cyclic Dependencies in Landmark Heuristics". Büchner, C., Keller, T., and Helmert, M. **Runner-Up for Best Student Paper Award at ICAPS 2021**
- "Learning General Optimal Policies with Graph Neural Networks: Expressive Power, Transparency, and Limits." Ståhlberg, S.; Bonet, B.; and Geffner, H. **Best Paper Award at ICAPS 2022**
- "Cost Partitioning Heuristics for Stochastic Shortest Path Problems". Klößner, T., Pommerening, F., Keller, T., and Röger, G. **Runner-Up for Best Student Paper Award at ICAPS 2022**
- "Additive Pattern Database for Decoupled Search". Sievers, S., Gnad, D., and Torralba, A. **Best Paper Award at SoCS 2022**
- "Towards ILP-Based LTLf Passive Learning". Ielo, A., Law, M., Fionda, V., Ricca, F., De Giacomo, G., and Russo, A. **Best Paper Award at ILP 2023**
- "Landmark Progression in Heuristic Search". Büchner, C., Keller, T., Eriksson, S., and Helmert, M. **Best Paper Award at ICAPS 2023**
- "Grounding Planning Tasks Using Tree Decompositions and Iterated Solving". Corrêa, A. B., Hecher, M., Helmert, M., Longo, D. M., Pommerening, F., and Woltran, S. **Runner-Up for Best Student Paper Award at ICAPS 2023**
- "Planning for Temporally Extended Goals in Pure-Past Linear Temporal Logic". Bonassi, L., De Giacomo, G., Favorito, M., Fuggitti, F., Gerevini A. E. and Scala, E. **Best Student Paper Award at ICAPS 2023.**
- "Integrating Planning and Learning for Agents Acting in Unknown Environments". Lamanna, L. **Best 2023 AI doctoral dissertation award (Marco Cadoli prize) - awarded by AI*IA - the Italian Association for Artificial Intelligence**