Project: **Kontiplan**

Towards Spacecraft On-board Autonomy

Contract No.: 50 RA 1220

Title: Technical Note, Final Report

Date: 31.12.2014

ABSTRACT

This technical note summarizes the application results and the scientific contributions of the project Kontiplan carried out under DLR contract by the Airbus DS GmbH in Friedrichshafen and the Institut für Informatik of the University Freiburg.

The overall goal of Kontiplan is to analyse and evaluate the application of elements from the field of artificial intelligence in different spaceflight scenarios. The elements in focus are action-based planning systems, so called action planners, which generate optimal solutions within decision trees. The following space scenarios

- In a **lander scenario** save landing sides shall be approached autonomously. For this, the planner considers dynamics, on-board resources and possible actions of the lander.
- In a **rendezvous & docking scenario** adequate approach strategies between two spacecraft shall be generated by automated planning.
- In an **Agile Earth observation scenario** for mission planning. The observation of different patches considering spacecraft motion capabilities and weather forecast is a challenging planning problem for which probabilistic planning algorithms shall be applied.

A major result of Kontiplan is the "Domain Predictive Control" approach that builds the bridge between dynamic (hybrid) systems and action planning. This approach allows the application of an action planner to control hybrid systems. Overall, out of the considered application scenarios, automated planning for agile Earth observation is the one that is best suited for applying an action planner.

TABLE OF CONTENTS

1 Project Motivation and Overview

This chapter gives the motivation for the project Kontiplan and provides an overview of the considered application scenarios (Lander, Rendezvous & Docking, Agile Earth Observation) and related scientific work.

The overall goal of Kontiplan is to analyse and evaluate the application of elements from the field of artificial intelligence in different spaceflight scenarios. The elements in focus are action-based planning systems, so called action planners, which generate optimal solutions within decision trees. Especially in time-critical mission phases without ground contact, such planning techniques shall increase the robustness of the spacecraft against uncertainties and unexpected events. In the future, by applying such algorithms on board in real-time, even complex decisions shall be made by the spacecraft itself and, thus, its autonomy level be increased.

The project Kontiplan focus' on the following space scenarios

- In a **lander scenario** save landing sides shall be approached autonomously. For this, the planner considers dynamics, on-board resources and possible actions of the lander.
- In a **rendezvous & docking scenario** adequate approach strategies between two spacecraft shall be generated by automated planning.
- In an **Agile Earth observation scenario** for mission planning. The observation of different patches considering spacecraft motion capabilities and weather forecast is a challenging planning problem for which probabilistic planning algorithms shall be applied.

However, the methods developed in Kontiplan are not limited to these scenarios. And from the enhancement of the planning algorithms, their application to a wider range of scenarios can be expected.

Lander Scenario

A lander scenario is analysed. The consideration of varying information content and quality about its environment during landing approach is required to find an adequate landing area by avoiding hazardous zones (e.g. craters, shadow). The adaptation of the lander trajectory is done autonomously due to absence of real-time communication links to the ground. The identification of hazardous zones is flagged by a separate system. Goal is to safely land on the surface of an unknown celestial body. By using planning algorithms, the landing trajectory shall be determined in a flexible sense and shall consider new informations collected during the landing procedure. Typical constraints to be taken into account are maximum landing speed, erroneous measurements, and limited on-board resources such as propellant mass.

Challenges: State-of-the art planning systems need to be enhanced and a method to be developed to translate lander dynamics into planning language. From a planning point-of-view, the big challenge is in the state-space structure: mainly numerical variables which, in addition, can have a huge number of different assignments. Another challenge is the required lander model complexity and, hence, the resulting computational load to generate plans. But a certain model complexity which is close to reality is necessary to generate realistic plans. One possible solution followed here is to outsource complex computations from the planning core algorithm into the so-called semantic attachments.

Rendezvous & Docking Scenario

A rendezvous & docking scenario is analysed. Goal is to autonomously dock a chaser spacecraft onto a target spacecraft which are not necessarily sharing the same orbit. For this, an approach strategy is required that shall be generated by planning. Given the relative dynamics of chaser and target spacecraft in form of the Hill equations, propellant-saving approaching strategies are to be searched for using predefined open- and closed-loop manoeuvres in the planning process.

Challenges: State-of-the art planning systems need to be enhanced and a method to be developed to translate lander dynamics into planning language. Like in the lander scenario, from a planning point-of-view, the big challenge is in the state-space structure: mainly numerical variables which, in addition, can have a huge number of different assignments. One possible solution followed here is to outsource complex computations from the planning core algorithm into the so-called semantic attachments. Another challenge is the uncertainty in case of an uncooperative target which explicitly calls for the use of probabilistic planning techniques.

Agile Earth Observation Scenario

An agile Earth observation scenario is analysed. Goal of the spacecraft is to observe a number of user/customer-defined patches on the Earth surface. The definition of each patch contains requirements on the spacecraft motion due to specific instrument properties. E.g. a required scan velocity on ground due to readout method of the instrument leads to spacecraft angular rate requirements. Both, optical and radar instruments are taken into account. Planning algorithms are used to bring these patches into an order that can be realized considering spacecraft agility constraints but, despite of that, maximises the information output for the customer. Agility constraints exist primarily due to actuator constraints. The application of planning algorithms shall fit into Astrium's reference process of agile Earth observation analysis.

Challenges: A simplification of the nonlinear satellite motion is required to make the problem solvable by the planning algorithm. Available on-board resources such as data memory but also instrument data rate, ground station visibility and station-specific downlink bandwidth are to be considered. In addition, varying patch illumination situations (day, dawn/dusk, night) and uncertain weather conditions (e.g. probabilistic cloud distribution) can limit the usage of some instruments and is to be taken into account.

Control of Hybrid Systems

The spacecraft in the scenarios introduced above can be handled are hybrid switched systems. They are hybrid systems because they contain both time-continuous numerical states (e.g. position, velocity, and attitude) and logical states (e.g. flags indicating actuator or sensor availability). And they are switched systems since the configuration of sensors, actuators and controllers can be actively changed during the mission. In order to apply planning algorithms to the scenarios, an appropriate method to control hybrid system is developed (Domain Predictive Control).

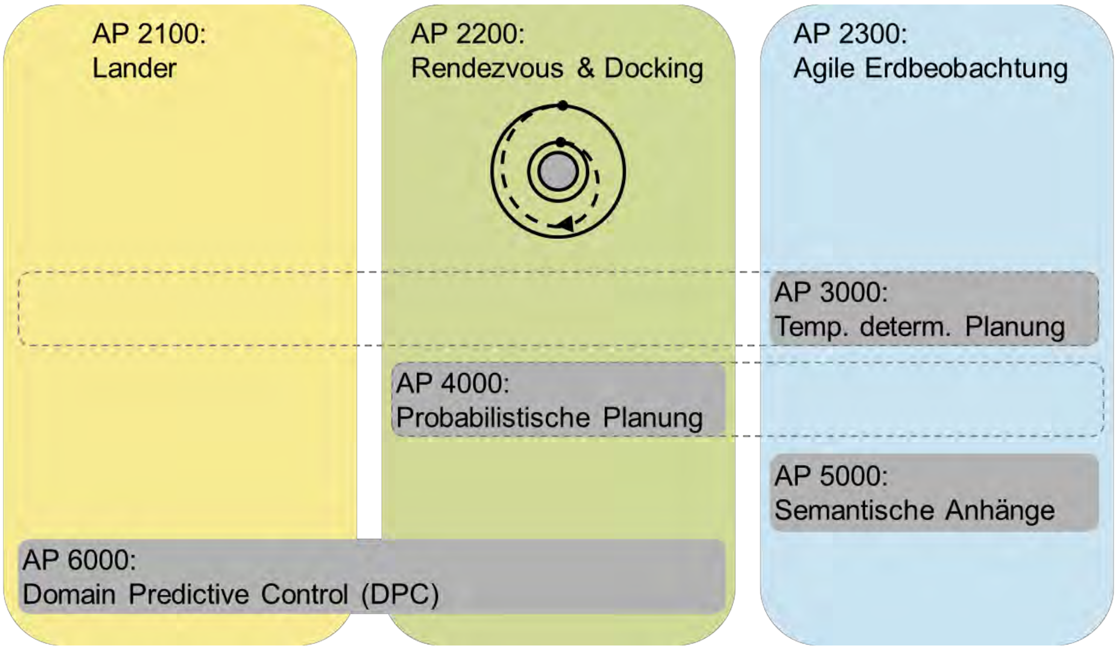

Figure 1-1: Application scenarios and related required scientific work topics in project Kontiplan.

2 Application Results

This chapter provides an overview of the currents results with focus on the space applications (1) Lander, (2) Rendezvous & Docking, and (3) Agile Earth-Observation.

2.1 Lander

In order to study the feasibility of applying planning methods to a lander mission, a prototype simulation was developed. Goal was to use planning for adapting the landing trajectory of a lander and taking unexpected events such as suddenly appearing craters or shadow fields (hazard avoidance) or failures (e.g. thruster) and resulting dynamical constrains into account. For this, the thrust vector alignments and related manoeuvres generated by the on-board propulsion system were set-up as actions for the planning algorithms (see Figure 2-2). Using these manoeuvers, the planner generated action sequences (plans) which guided the lander to the ground. Goal was to reach a landing zone with lander-specific maximum impact velocity (typically 1.5-2.0 m/s).

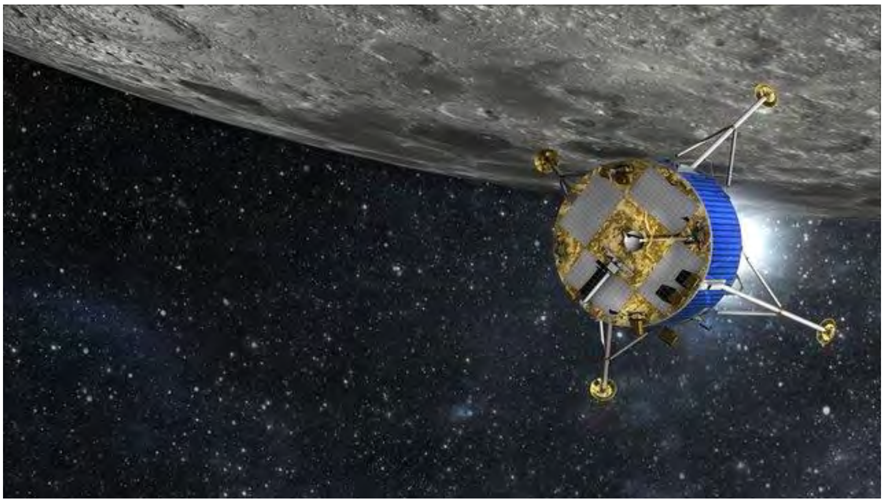

Figure 2-1: Illustration of a lunar lander.

In order to embed the planning system into the prototype simulation, all components were linked within an appropriate architecture and the Domain Predictive control, developed within Kontiplan, was applied. The case study focused on short thrust manoeuvres of 2 s duration and slews of 5 s duration to change attitude between A, B, C and D. It was shown that the interaction between the Matlab simulation and the planning system Temporal Fast Downward (TFD) is feasible over clearly defined interfaces. Trajectory planning and replanning was sucessfully demonstrated as shown in Figure 2-3.

In conclusion, the application of planning in order to adapt control signals and trajectories of landers was analyzed. The feasibility of choosing appropriate control signals by planing-based prediction and evaluation of future system states could be demonstrated by simulation. For this, the possibility of handling conditions and dependencies containing both numerical and logical states is one of the most gainful aspects. Experiences with Domain Predictive Control show that the development of a well balanced planning domain is the most challenging point of this approach. This is due to the high computational effort of planning with complex domains.

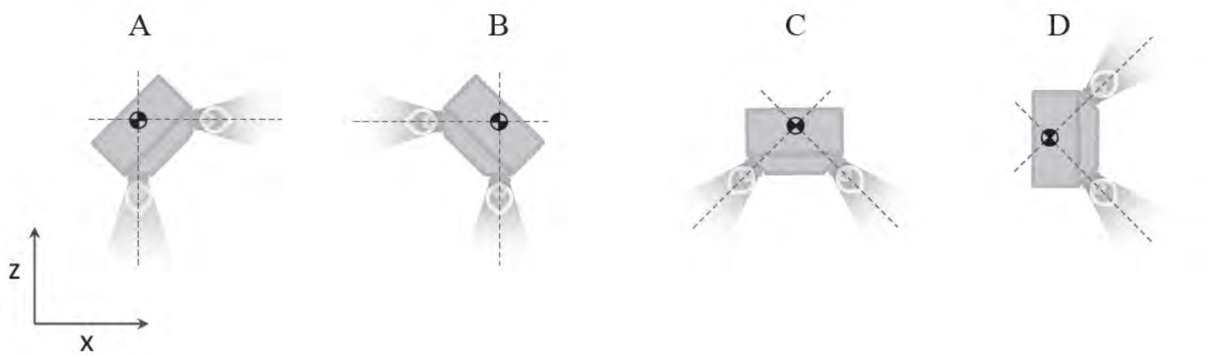

Figure 2-2: Considered lander attitudes. Slew manoeuver is required between attitudes A, B, C and D.

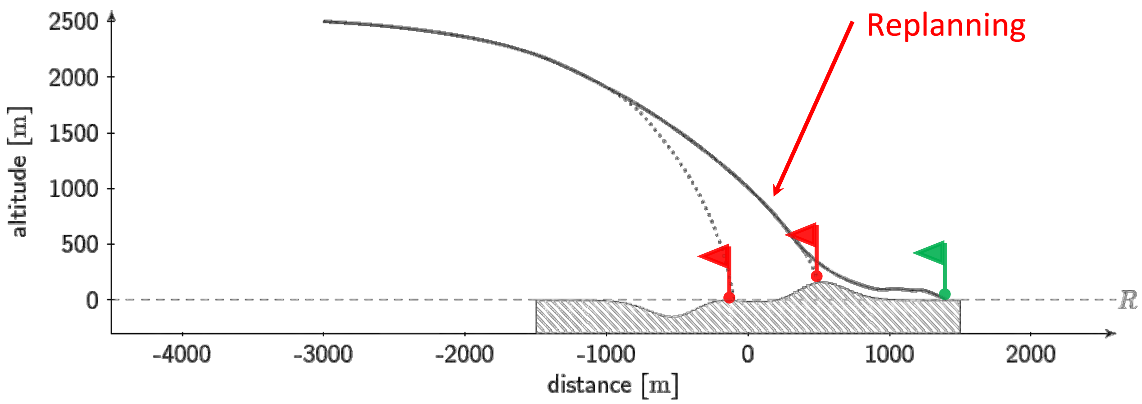

Figure 2-3: Simulation results of spacecraft trajectory using Domain Predictive Control method for successful landing despite late landing site relocation.

2.2 Rendezvous & Docking

In order to study the feasibility of applying planning methods to a rendezvous and docking mission, a prototype simulation was developed. Goal was to use planning in order to generate an approach strategy for the chaser spacecraft to the target spacecraft. For this, classical orbital manoeuvres based on the Hill-equations (also known as Clohessy-Wiltshire equations) were used as actions for the planner. Using these actions, the planner generated sequences which guided the chaser into the vicinity of the target spacecraft located into the origin of the Hill system (here: 500 x 50 m^2 box within orbital plane).

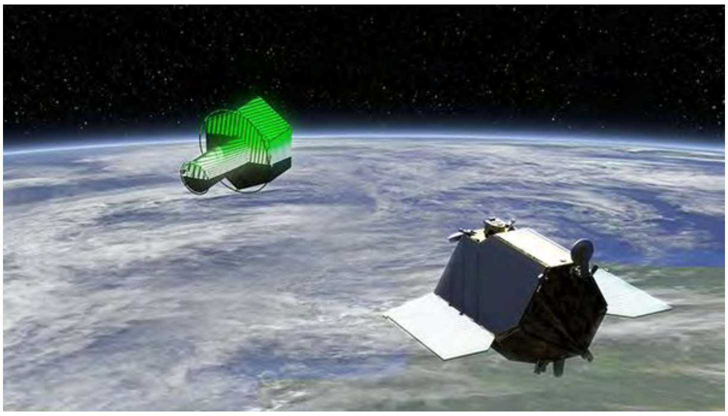

Figure 2-4: Approaching an uncooperative target.

As in the lander study in section 2.1, all components were linked within an appropriate architecture in order to link the planning system with the prototype simulation. Short thrust manoeuvres were considered to raise apogee and/or perigee. The full planning concept and related actions are shown in Figure 2-5. It was shown that the interaction between the Matlab simulation and the planning system Temporal Fast Downward (TFD) is feasible over clearly defined interfaces.

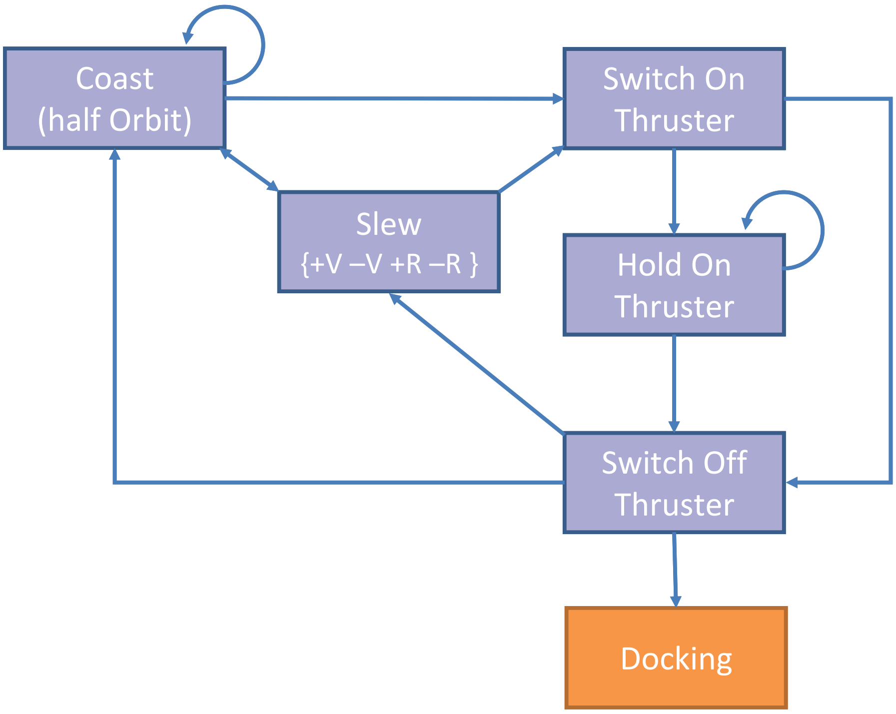

Figure 2-5: Planning concept for approach strategy generation in rendezvous & docking scenario.

The robustness of the generated plan was is special focus of this application. Therefore, the privious (nominal) planning was enhanced by taking uncertainties, as they always occur in reality, during the planning process into account. As this domain can be discretized and linearized straight-forwardly, Domain Predictive Control is well suited to generate manoeuver-based approach strategies in principle. However, because actions are relatively long, small uncertainties in the initial state typically accumulate to large errors in the goal state.

Both uncertain initial conditions and noisy thrusters and sensors can affect the final position such that the found plan leads to a potential crash with the other spacecraft. The orbit manoeuvers are planned within the Nadir frame with origin at the goal position, moving on a circular orbit around Earth.

In the considered case study, a simplified domain model is used. The thruster provides a force in flight direction and can either be switched on or off. A controller is used to avoid movements of the spacecraft in radial direction while thrusting in velocity direction. Additional possibilities of thrusting in ±radial direction resulting in circular fly-around or negative velocity direction thrusts are not considered for simplicity. The result of the planning process without propagation of the error state covariance using Temporal Fast Downward is a plan that hits the designated goal position, see Figure 2-6A. However, uncertainties of the initial states and disturbances during plan execution let the spacecraft possibly crash with the target spacecraft. Propagation of the covariance matrix during the planning process yields a different plan. It reaches the goal position and makes the 3σ-confidence ellipse matching the desired orbit zone, see Figure 2-6B. Allowing for one reference tracking phase with active error controller leads to another plan. The planned reference tracking phase actively damps out the error states and leads to a final state satisfying the goal conditions, see Figure 2-6B.

In conclusion, all plans were computed in a few seconds and ensure that the goal position is reached with the predefined confidence by either reordering actions intelligently or by performing reference tracking. In contrast, planning without reasoning about errors and uncertainties does not give any information about the probability to reach the goal position.

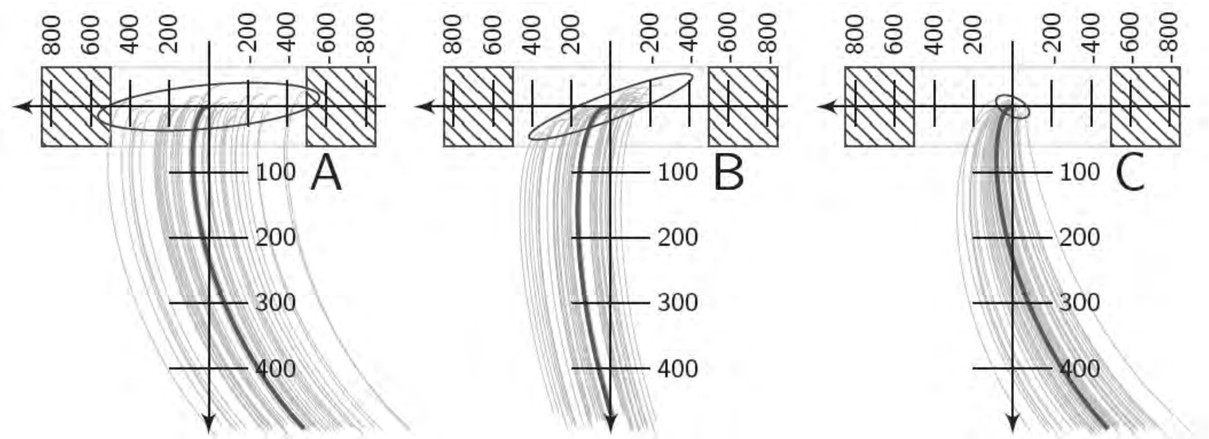

Figure 2-6: Approach trajectories close to the goal positions at (0,0) and final 3σ-confidence ellipses: (A) without and (B) with covariance propagation, (C) with planned reference tracking phase.

In addition to applying the planner Temporal Fast Downward to the rendezvous and docking scenario, also the feasibility of using the probabilistic planner PROST was analysed. A simplified dynamical model was considered describing the spacecraft position as 2-dimensional coordinates. The considered actions were (1) thruster firing for velocity increase in one direction and (2) thruster firing for velocity decrease in one direction. As in reality, all actions could fail with a known probability. In addition, collision of chaser and target must be avoided at all costs. The planning results are shown in Figure 2-7. One might expect the trajectory from the “Start” to the “Target area” being a straight line to be the best solution. It clearly is the shortest but in case of thruster failure close to the target along x-direction, a collision could not be avoided. Therefore the solution generated by PROST presents a safer, collision-avoiding approach.

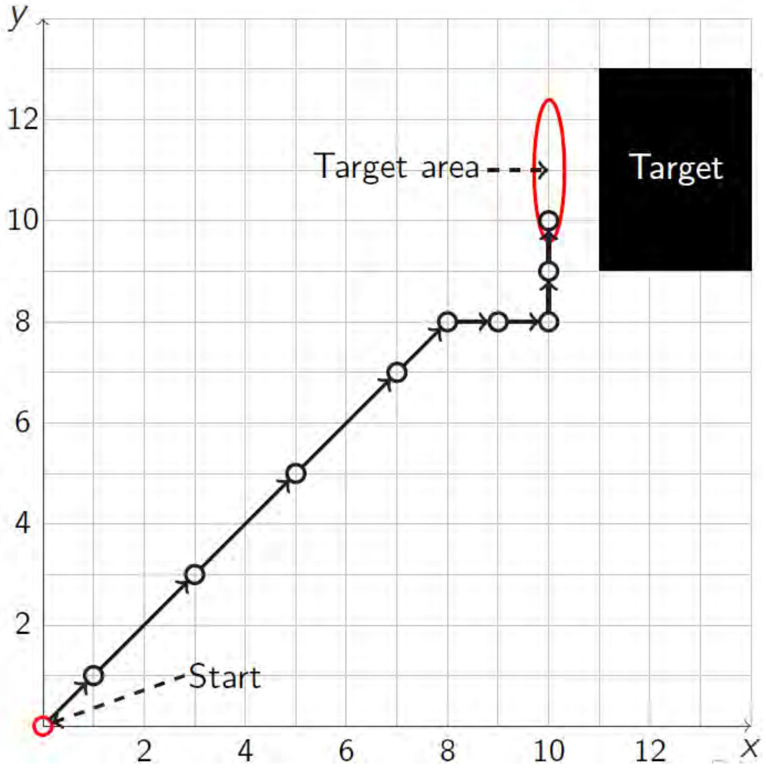

Figure 2-7: Exemplary approach strategy considering thruster failure within the planning process.

2.3 Agile Earth-Observation

In order to study the feasibility of applying planning methods to an agile Earth observation mission, the related Astrium analysis process using the internal tool “Asset” was enhanced by a “Planner” step as shown in Figure 2-8. Goal was to use planning in order to generate an optimal order of areas on Earth, called “patch”, to be scanned by the spacecraft.

In agile observation missions the observation is done with instruments rigidly mounted on the satellite's body. The resolution of the images can be increased by larger sensors or the use of instruments with smaller field-of-view yielding a reduced coverage of the Earth's surface during each revolution of the satellite in its orbit. In such missions it is aspired to scan designated observation sites rather than to achieve global coverage of the Earth's surface. This requires the instrument or the whole satellite to slew all axes like in the Pleiades, Eros or WorldView missions. Instead of taking pictures with matrix sensors, here it was focused on instruments which continuously scan an observation site using time-delayed integration. Therefore, a specific relative motion of the instrument's line-of-sight with respect to the observation site was realized. The result is an observation strip on the Earth's surface, a patch. An observation scenario is defined by a large set of patches arbitrarily distributed on Earth with predefined priorities which shall reect the customer's need.

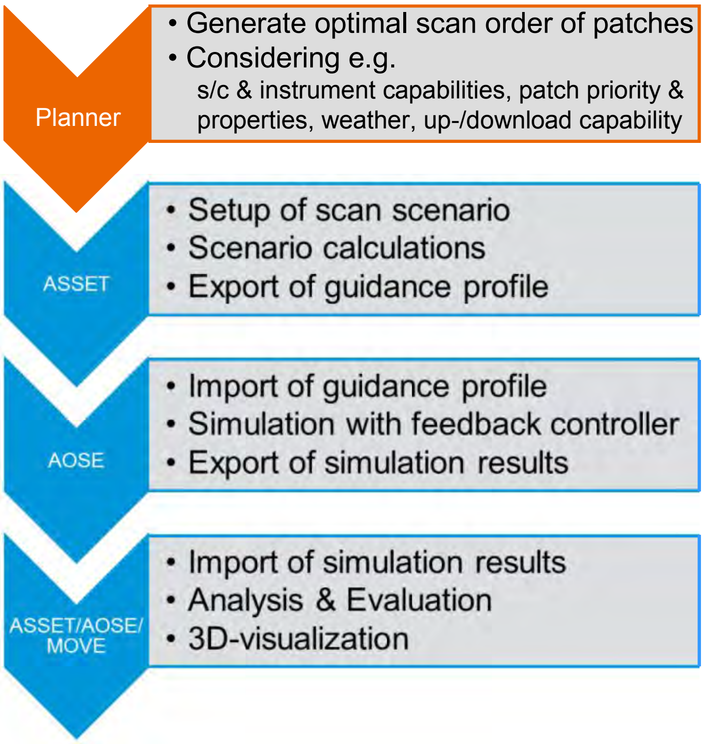

Figure 2-8: Reference process of agile Earth observation analysis including the new “Planner” step.

The goal of the analysis was to find suitable sequences of observation patches yielding a feasible sequence of slew maneuvers, considering the satellite's orbital motion, its attitude and angular rate as well as its torque capability in realistic scenarios. Such a sequence (plan) was generated by a planning algorithm.

Instrument alignment and the required scan velocity pose constraints within the planning process. The feasibility of slews between two successive scans depends on the satellite's attitude, angular rate and position and is varying in time. Any decision to scan a certain patch at a certain time may affect the feasibility of future scan maneuvers. This makes the problem dificult to solve in case of larger sets of observation patches.

An algorithm was developed which considers the priority of the patches and is not based on a simplified dynamic, since it directly uses the functionality of the Asset toolbox, which calculates the whole satellite guidance including optimized slew maneuvers for a given patch sequence. This new planning algorithm neglects observable patches only, if their observation inhibits the scan of patches with higher priority. For this the Earth observation task was reformulated into a planning problem which is solvable by the proposed planning algorithm. In order to reduce the complexity of the planning problem before starting with planning, as many infeasible sequences as possible were canceled out during a preprocessing phase (planning horizon, visibility, overlap, transition).

For evaluation of the developed algorithm, an examplary agile Earth observation example was generated. For this, 5000 randomly distributed observation patches of length from 100 km to 600 km and arbitrary orientation were generated, see Figure 2-9. This scenario was highly over-specified and the goal was to find a plan with duration of one orbit period which includes as many patches as possible considering the priority conditions discussed before. A satellite with moment of inertia of 500 kgm^2 about each axis and a maximum available torque of 1 Nm was considered. The body-fixed instrument provides scans of acceptable quality if both the roll angle and the pitch angle are below 35 deg with respect to Nadir. The maximally acceptable angular rate was 10 deg/s about each axis. The results of this example for first planning horizon are shown in Figure 2-10 and for a complete orbit in Figure 2-10. They confirm the feasibility of the developed approach.

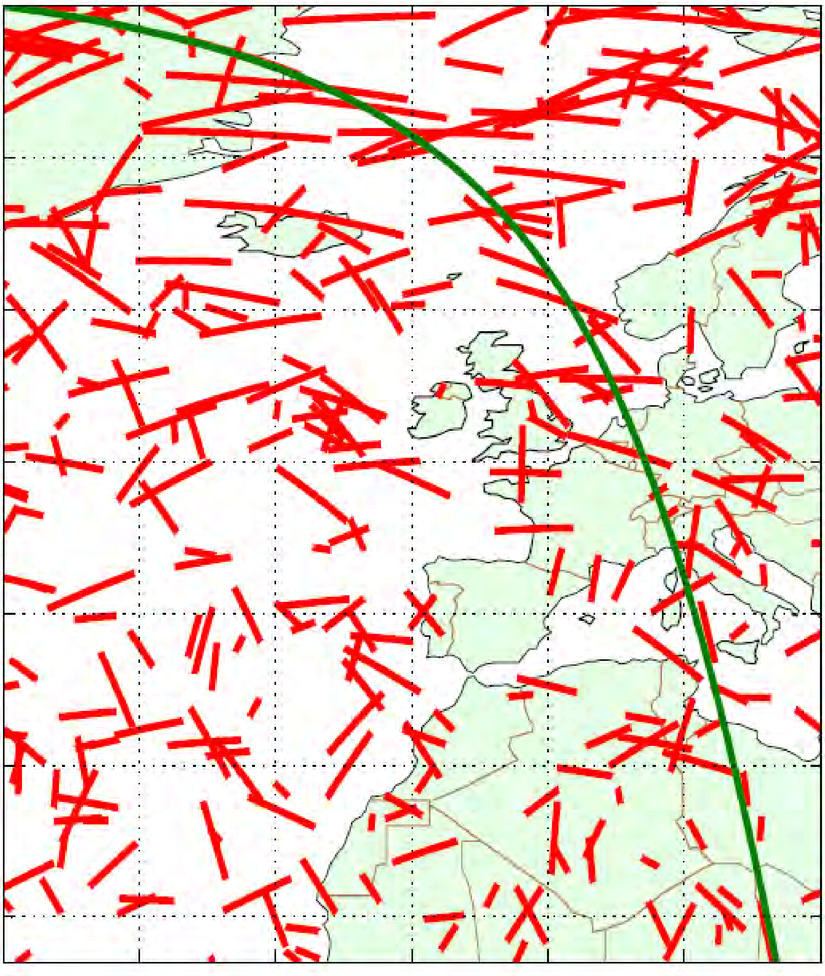

Figure 2-9: Random scenario with view over Europe. Observation patches are depicted as red stripes while the ground track of the satellite shown in green.

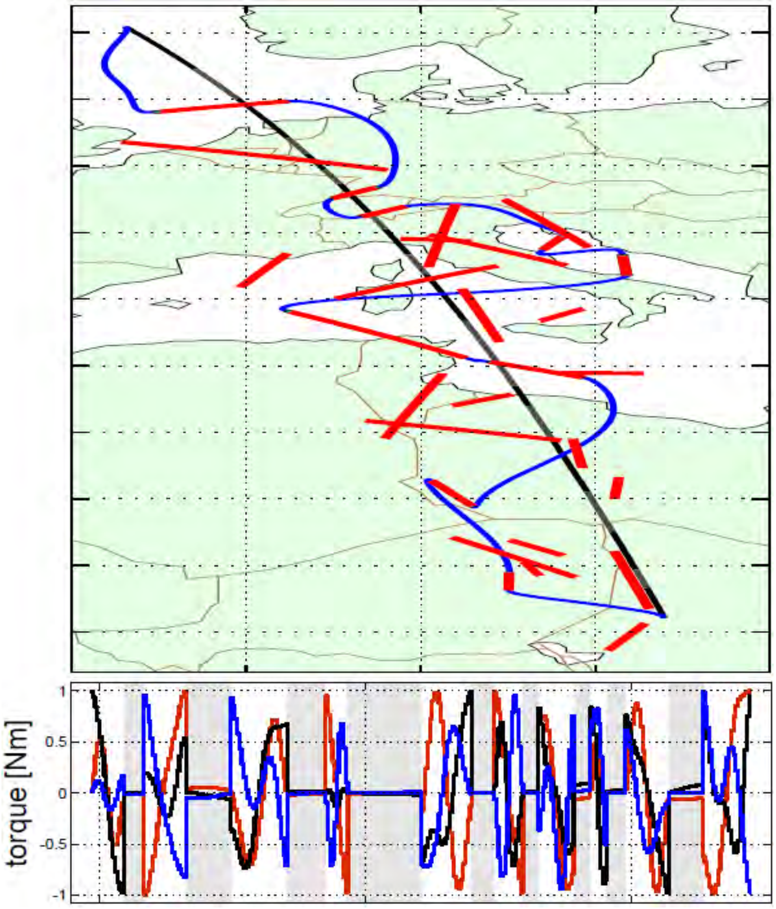

Figure 2-10: Optimal sequence for the first planning horizon with visible patches in red, ground track in black and intersection between line of sight and Earth's surface in blue. The ground track is shown in grey while scanning. The torque demand is within the specified limits of 1 Nm in each axis.

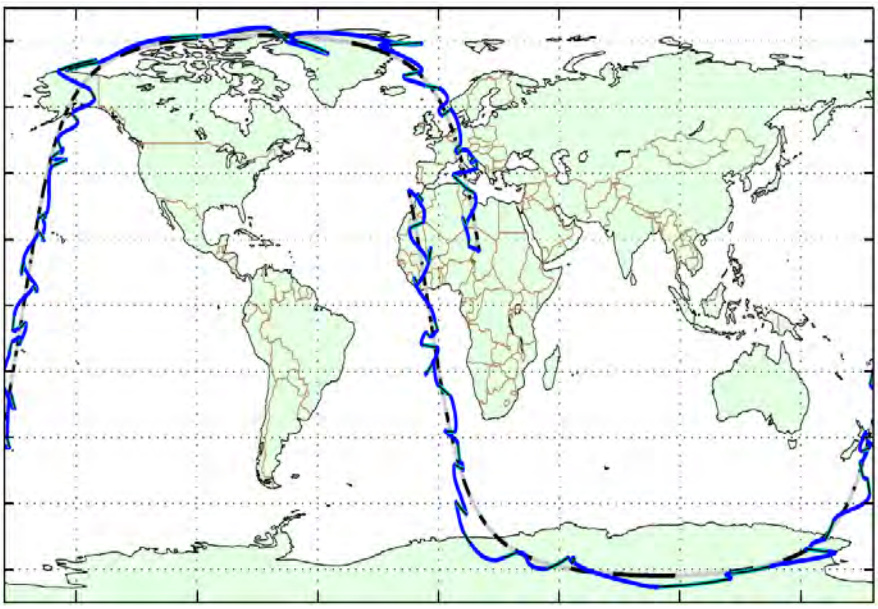

Figure 2-11: The final plan over one orbit. It corresponds to the guidance profile of the satellite over one orbit connecting 61 observation sites with the line of sight of the instrument on Earth, shown in blue. The ground track of the satellite is shown in black during slews and in grey while scanning.

2.4 Summary and Conclusion

From the space-industry point of view, before Kontiplan the field of research in action planning was primarily focused on artificial "applications". Real world complexity and constraints of (dynamic) systems such as numerical states or nonlinear motion, which are common daily aspects of industrial work, were barely or even not considered at all. Instead of real world applications, the focus was/is more on standardized artificial/theoretical action planning problems and competitions of and within the planning community. With the project Kontiplan the chance was provided to apply the theoretical methods to space applications, to systems with discrete and continuous (numerical) states, thus hybrid systems.

A major result of Kontiplan is the "Domain Predictive Control" approach that builds the bridge between dynamic (hybrid) systems and action planning. This approach allows the application of an action planner to control hybrid systems. Regarding applications, the major limitation is the maximum possible number of actions to ensure that a plan can be generated (computing time!). Thus, currently only application to low-frequent actuations seam feasible (e.g. in-orbit approach phase).

Overall, out of the considered application scenarios, automated planning for agile Earth observation is the one that is best suited for applying an action planner.

2.5 Future Work

Before applying the Domain Predictive Control approach in a real space mission, it must be challenged by established solutions. A solid and comprehensive comparison of state-of-the-art solutions from AOCS/GNC should be carried out.

In order to better fit current and upcoming Earth observation requests by Airbus DS' customers, the developed automated planning tool should be extended to multi-satellite scenarios of different payloads (optical, radar) and to interaction with other systems such as ground stations and UAVs.

3 Scientific Contribution

This chapter provides the scientific contributions of the project Kontiplan.

The scientific contribution cover the following topics

- Prost: Probabilistic Planning Based on UCT
- A Planning Based Framework for Controlling Hybrid Systems
- Planning-based Autonomous Lander Control
- Preferring Properly: Increasing Coverage while Maintaining Quality in Anytime Temporal Planning
- Planning for Agile Earth Observation Satellites
- Stronger Abstraction Heuristics Through Perimeter Search
- Trial-based Heuristic Tree Search for Finite Horizon MDPs
- Domain Predictive Control Under Uncertain Numerical State Information
- Automated Planning for Earth Observation Spacecraft under Attitude Dynamical Constraints
- Trial-based Heuristic Tree Search for Finite Horizon MDPs
- Symbolic Domain Predictive Control
- Balancing Exploration and Exploitation in Classical Planning
- Past, Present, and Future: An Optimal Online Algorithm for Single-Player GDL-II Games
- An Experimental Comparison of Classical, FOND and Probabilistic Planning
- Automated Planning for Agile Earth Observation

In the following, all papers are provided including the links to the related Kontiplan work packages.

Thomas Keller und Patrick Eyerich

Prost: Probabilistic Planning Based on UCT

ICAPS 2012

Abstract

We present Prost, a probabilistic planning system that is based on the UCT algorithm by Kocsis and Szepesvari (2006), which has been applied successfully to many areas of planning and acting under uncertainty. The objective of this paper is to show the application of UCT to domain-independent probabilistic planning, an area it had not been applied to before. We furthermore present several enhancements to the algorithm, including a method that is able to drastically reduce the branching factor by identifying superfluous actions. We show how search depth limitation leads to a more thoroughly investigated search space in parts that are influential on the quality of a policy, and present a sound and polynomially computable detection of reward locks, states that correspond to, e.g., dead ends or goals. We describe a general Q-value initialization for unvisited nodes in the search tree that circumvents the initial random walks inherent to UCT, and leads to a faster convergence on average. We demonstrate the significant influence of the enhancements by providing a comparison on the IPPC benchmark domains.

Zugeordnete Arbeitspakete

AP 4200: Weiterentwicklung der Initialisierung sowie der automatischen Sackgassen- und Zielerkennung

Johannes Löhr, Patrick Eyerich, Thomas Keller, and Bernhard Nebel

A Planning Based Framework for Controlling Hybrid Systems

Published on the 22rd Conference on Automated Planning and Scheduling, ICAPS 2012, Atibaia, Brazil

Abstract

The control of dynamic systems, which aims to minimize the deviation of state variables from reference values in a continuous state space, is a central domain of cybernetics and control theory. The objective of action planning is to find feasible state trajectories in a discrete state space from an initial state to a state satisfying the goal conditions, which in principle addresses the same issue on a more abstract level. We combine these approaches to switch between dynamic system characteristics on the fly, and to generate control input sequences that affect both discrete and continuous state variables. Our approach (called Domain Predictive Control) is applicable to hybrid systems with linear dynamics and discretizable inputs.

Zugeordnete Arbeitspakete

AP 6000: Steuerung Hybrider Systeme mittels Domain Predictive Control
AP 6100: Systematische Methode zum Einsatz von Planungstechniken zum steuern dynamischer Systeme

Johannes Löhr, Bernhard Nebel, and Stefan Winkler

Planning-based Autonomous Lander Control

Published on the Astrodynamics Specialist Conference, AIAA/AAS 2012, Minneapolis, USA

Abstract

Safe landing of spacecraft on extraterrestrial surfaces implies a number of challenges. The main issue is to precisely initiate coasting, braking and landing maneuvers to safely land at a desired landing zone. Meanwhile, the increasing information level about the landing environment has to be processed and the landing trajectory eventually adapted in order to avoid hazardous situations. In this paper these time critical tasks are performed by Domain Predictive Control. It has been developed to guide dynamic systems into desired goal states by flexibly reordering atomic actions using planning algorithms from artificial intelligence. Here, the method is applied to autonomously adapt control commands and associated landing trajectories with respect to the changing environmental knowledge. Simulation results show the feasibility of this new approach and reveal issues which should be subject to future research.

Zugeordnete Arbeitspakete

AP 2100: Lander
AP 2110: Konzepterstellung
AP 2120: Modellierung
AP 2130: Simulation
AP 6000: Steuerung hybrider Systeme mittels Domain Predictive Control

Patrick Eyerich

Preferring Properly: Increasing Coverage while Maintaining Quality in Anytime Temporal Planning

ECAI 2012

Abstract

Temporal Fast Downward (TFD) is a successful temporal planning system that is capable of dealing with numerical values. Rather than decoupling action selection from scheduling, it searches directly in the space of time-stamped states, an approach that has shown to produce plans of high quality at the price of coverage. To increase coverage, TFD incorporates deferred evaluation and preferred operators, two search techniques that usually decrease the number of heuristic calculations by a large amount. However, the current definition of preferred operators offers only limited guidance in problems where heuristic estimates are weak or where subgoals require the execution of mutex operators. In this paper, we present novel methods for refinement of this definition and show how to combine the diverse strengths of different sets of preferred operators using a restarting procedure incorporated into a multi-queue best-first search. These techniques improve TFD's coverage drastically and preserve the average solution quality, leading to a system that solves more problems than each of the competitors of the temporal satisficing track of IPC 2011 and clearly outperforms all of them in terms of IPC score.

Zugeordnete Arbeitspakete

AP 3200: Bevorzugte Operatoren

Johannes Aldinger and Johannes Löhr

Planning for Agile Earth Observation Satellites

ICAPS-2013 Workshop on Planning in Continuous Domains (PCD)

Abstract

Agile Earth observation satellites are satellites orbiting Earth with the purpose to gather information of the Earth's surface by slewing the satellite toward regions of interest. Constraints arise not only from dynamical and kinematic aspects of the satellite and its sensors. Regions of interest change over time and bad weather can conceal important observation targets. This results in a constant need to replan the satellite's tasks and raises the desire to automatize this planning process. We consider the Earth observation problem with the help of the module extension of the numerical planning system Temporal Fast Downward. Complex satellite slew maneuvers are calculated within modules, while the planner selects and schedules the regions to be scanned. First results encourage deeper research in this area so that forthcoming satellite space missions can draw on automated planning to improve the performance of agile Earth observation tasks.

Zugeordnete Arbeitspakete

AP 2300: Szenario: Agile Erdbeobachtung
AP 5000: Objekt-orientierte Planung mit semantischen Anhängen
AP 5100: Fertigstellung von Basisarbeiten
AP 5300: Anpassung von Heursitiken

Patrick Eyerich und Malte Helmert

Stronger Abstraction Heuristics Through Perimeter Search

ICAPS 2013

Abstract

Perimeter search is a bidirectional search algorithm consisting of two phases. In the first phase, a limited regression search computes the perimeter, a region which must necessarily be passed in every solution. In the second phase, a heuristic forward search finds an optimal plan from the initial state to the perimeter.
The drawback of perimeter search is the need to compute heuristic estimates towards every state on the perimeter in the forward phase. We show that this limitation can be effectively overcome when using pattern database (PDB) heuristics in the forward phase.
The combination of perimeter search and PDB heuristics has been considered previously by Felner and Ofek for solving combinatorial puzzles. They claimed that, based on theoretical considerations and experimental evidence, the use of perimeter search in this context offers "limited or no benefits". Our theoretical and experimental results show that this assessment should be revisited.

Zugeordnete Arbeitspakete

AP 3300: Alternative Suchalgorithmen im Kontext Verzögerter Evaluierung

Thomas Keller und Malte Helmert

Trial-based Heuristic Tree Search for Finite Horizon MDPs

ICAPS 2013 & RLDM 2013

Abstract

Dynamic programming is a well-known approach for solving MDPs. In large state spaces, asynchronous versions like Real-Time Dynamic Programming have been applied successfully. If unfolded into equivalent trees, Monte-Carlo Tree Search algorithms are a valid alternative. UCT, the most popular representative, obtains good anytime behavior by guiding the search towards promising areas of the search tree. The Heuristic Search algorithm AO* finds optimal solutions for MDPs that can be represented as acyclic AND/OR graphs.
We introduce a common framework, Trial-based Heuristic Tree Search, that subsumes these approaches and distinguishes them based on five ingredients: heuristic function, backup function, action selection, outcome selection, and trial length. Using this framework, we describe three new algorithms which mix these ingredients in novel ways in an attempt to combine their different strengths. Our evaluation shows that two of our algorithms not only provide superior theoretical properties to UCT, but also outperform state-of-the-art approaches experimentally.

Zugeordnete Arbeitspakete

AP 4100: Analyse und Entwicklung von Baumsuchalgorithmen

Johannes Löhr, Patrick Eyerich, Stefan Winkler, and Bernhard Nebel

Domain Predictive Control Under Uncertain Numerical State Information

Published on the 23rd Conference on Automated Planning and Scheduling, ICAPS 2013, Rome, Italy

Abstract

In planning, hybrid system states consisting of logical and numerical variables are usually assumed to be completely known. In particular, for numerical state variables full knowledge of their exact values is assumed. However, in real world applications states are results of noisy measurements and imperfect actuators. Therefore, a planned sequence of state transitions might fail to lead a hybrid system to the desired goal. We show how to propagate and reason about uncertain state information directly in the planning process, enabling hybrid systems to find plans that satisfy numerical goals with predefined confidence.

Zugeordnete Arbeitspakete

AP 6000: Steuerung Hybrider Systeme mittels Domain Predictive Control
AP 6300: Planen mit Unsicherheiten in Systemzuständen
AP 6400: Evaluation DPC
AP 2200: Relativdynamik zweier Raumfahrzeuge als Anwendungsszenario

Johannes Löhr, Johannes Aldinger, Stefan Winkler, and Georg Willich

Automated Planning for Earth Observation Spacecraft under Attitude Dynamical Constraints

Veröffenlicht auf dem Deutschen Luft- und Raumfahrtkongress, DGLR 2013, Stuttgart, Deutschland

Abstract

Agile Earth observation missions continuously require a large amount of planning during the spacecraft's observations. Beside priorities of the observation sites, especially the agility constraints of the satellite are important to be taken into account during the planning process. This is due to the body-fixed instrument's line of sight, requiring the whole satellite to point to the observation sites while scanning. Scanning a sequence of observation sites leads to complex slew maneuvers which must not exceed the satellite's actuator capacities, attitude constraints or maximum angular rates. Additionally, the regions of interest may change over time, making it necessary to adapt and optimize the observation sequence continuously. An automated process is required to efficiently handle this task. We present a planning algorithm to sequence an arbitrarily distributed set of observation patches to a feasible observation plan, considering priority criteria of the observation sites and agility constraints of the satellite.

Zugeordnete Arbeitspakete

AP 2300: Agile Erdbeobachtung
AP 2130: Konzepterstellung
AP 2130: Modellierung
AP 2130: Simulation

Thomas Keller and Malte Helmert

Trial-based Heuristic Tree Search for Finite Horizon MDPs

Conference on Reinforcement Learning and Decision Making 2013

Abstract

Dynamic programming is a well-known approach for solving MDPs. In large state spaces, asynchronous versions like Real-Time Dynamic Programming (RTDP) have been applied successfully. If unfolded into equivalent trees, Monte-Carlo Tree Search algorithms are a valid alternative. UCT, the most popular representative, obtains good anytime behavior by guiding the search towards promising areas of the search tree and supporting non-admissible heuristics. The global Heuristic Search algorithm AO* finds optimal solutions for MDPs that can be represented as acyclic AND/OR graphs. Despite the differences, these approaches actually have much in common. We present the Trial-based Heuristic Tree Search (THTS) framework that subsumes these approaches and distinguishes them based on only five ingredients: heuristic function, backup function, action selection, outcome selection, and trial length. We describe the ingredients that model RTDP, AO* and UCT within this framework, and use THTS to combine attributes of these algorithms step by step in order to derive novel algorithms with superior theoretical properties. We merge Full Bellman and Monte-Carlo backup functions to Partial Bellman backups, and gain a function that both allows partial updates and a procedure that labels states when they are solved. DP-UCT combines attributes and theoretical properties from RTDP and UCT even though it differs from the latter only in the used Partial Bellman backups. Our main algorithm, UCT* adds a limited trial length to DP-UCT to inherit the global search behavior of AO* , which ensures that parts of the state space that are closer to the root are investigated more thoroughly. The experimental evaluation shows that both DP-UCT and UCT* are not only superior to UCT, but also outperform P ROST , the winner of the International Probabilistic Planning Competition (IPPC) 2011 on the benchmarks of IPPC 2011.

Zugeordnete Arbeitspakete

AP 4100: Analyse und Entwicklung von Baumsuchalgorithmen

Johannes Löhr, Martin Wehrle, Maria Fox, and Bernhard Nebel

Symbolic Domain Predictive Control

**Published on the 28th Conference Atificial Intelligence,
AAAI 2014, Québec, Kanada**

Abstract

Planning-based methods to guide switched hybrid systems from an initial state into a desired goal region opens an interesting field for control. The idea of the Domain Predictive Control (DPC) approach is to generate input signals affecting both the numerical states and the modes of the system by stringing together atomic actions to a logically consistent plan. However, the existing DPC approach is restricted in the sense that a discrete and pre-defined input signal is required for each action. In this paper, we extend the approach to deal with symbolic states. This allows for the propagation of reachable regions of the state space emerging from actions with inputs that can be arbitrarily chosen within specified input bounds. This symbolic extension enables the applicability of DPC to systems with bounded inputs sets and increases its robustness due to the implicitly reduced search space. Moreover, precise numeric goal states instead of goal regions become reachable.

Zugeordnete Arbeitspakete

AP 6000: Steuerung hybrider Systeme mittels Domain Predicitve Control
AP 6100: Systematische Methodik zur symbolischen Steuerung hybrider Systeme
AP 6200: Stabilitätsaspekte
AP 6400: Evaluation
AP 2200: Relativdynamik zweier Raumfahrzeuge als Anwendungszenario

Tim Schulte und Thomas Keller

Balancing Exploration and Exploitation in Classical Planning

SOCS 2014

Abstract

Successful heuristic search planners for satisficing planning like FF or LAMA are usually based on one or more best first search techniques. Recent research has led to planners like Arvand, Roamer or Probe, where novel techniques like Monte-Carlo Random Walks extend the traditional exploitation-focused best first search by an exploration component. The UCT algorithm balances these contradictory incentives and has shown tremendous success in related areas of sequential decision making but has never been applied to classical planning yet. We make up for this shortcoming by applying the Trial-based Heuristic Tree Search framework to classical planning. We show how to model the best first search techniques Weighted A* and Greedy Best First Search with only three ingredients: action selection, initialization and backup function. Then we use THTS to derive four versions of the UCT algorithm that differ in the used backup functions. The experimental evaluation shows that our main algorithm, GreedyUCT, outperforms all other algorithms presented in this paper, both in terms of coverage and quality.

Zugeordnete Arbeitspakete

AP 3000: Temporale deterministische Planung in numerischen Zustandsräumen

Florian Geißer, Thomas Keller und Robert Mattmüller

Past, Present, and Future: An Optimal Online Algorithm for Single-Player GDL-II Games

ECAI 2014

Abstract

In General Game Playing, a player receives the rules of an unknown game and attempts to maximize his expected reward. Since 2011, the GDL-II rule language extension allows the formulation of nondeterministic and partially observable games. In this paper, we present an algorithm for such games, with a focus on the single-player case. Conceptually, at each stage, the proposed N ORNS algorithm distinguishes between the past, present and future steps of the game. More specifically, a belief state tree is used to simulate a potential past that leads to a present that is consistent with received observations. Unlike other related methods, our method is asymptotically optimal. Moreover, augmenting the belief state tree with iteratively improved probabilities speeds up the process over time significantly.
As this allows a true picture of the present, we additionally present an optimal version of the well-known UCT algorithm for partially observable single-player games. Instead of performing hindsight optimization on a simplified, fully observable tree, the true future is simulated on an action-observation tree that takes partial observability into account. The expected reward estimates of applicable actions converge towards the true expected rewards even for moves that are only used to gather information. We prove that our algorithm is asymptotically optimal for single-player games and POMDPs and support our claim with an empirical evaluation.

Zugeordnete Arbeitspakete

AP 4000: Probabilistische Planung

Andreas Hertle, Christian Dornhege, Thomas Keller, Robert Mattmüller, Manuela Ortlieb und Bernhard Nebel

An Experimental Comparison of Classical, FOND and Probabilistic Planning

KI 2014

Abstract

Domain-independent planning in general is broadly applicable to a wide range of tasks. Many formalisms exist that allow the description of different aspects of realistic problems. Which one to use is often no obvious choice, since a higher degree of expressiveness usually comes with an increased planning time and/or a decreased policy quality. Under the assumption that hard guarantees are not required, users are faced with a decision between multiple approaches. As a generic model we use a probabilistic description in the form of Markov Decision Processes (MDPs). We define abstracting translations into a classical planning formalism and fully observable nondeterministic planning. Our goal is to give insight into how state-of-the-art systems perform on different MDP planning domains.

Zugeordnete Arbeitspakete

AP 2300: Agile Erdbeobachtung
AP 4300: Unterstützung numerischer Variablen

Maja Puhle, Johannes Löhr, Stefan Winkler

Automated Planning for Agile Earth Observation

Abstract

Agile Earth observation satellites represent currently the best option to acquire requested images of the Earth surface during a short observation time performing fast attitude maneuvers between the observations. Thereby control moment gyroscopes are used as attitude actuators providing the demanded high control torques and enabling precision pointing and tracking simultaneously. Time delayed integration sensors are used to capture high resolution images. As a consequence, the maneuvers have to be adapted to these devices and it has to be taken into consideration that control moment gyroscopes can reach states where no torque output in specific directions is producible during their operation. These arising singularities have to be considered during the maneuver planning. Furthermore the selection and scheduling process becomes more challenging for agile satellites since the number of observable areas increases notably. In the present work the usage of CMGs for agile Earth observation satellites is investigated and the correct modeling of control moment gyroscope based maneuvers is integrated in a tool which is used for the feasibility analysis of predefined observation sequences. Thereby the run-time of the program is reduced which is essential to deal effectively with the exponentially increasing number of sequence feasibility tests. In the following a second tool that generates feasible observation schedules is presented and the undertaken changes and extensions of the program are pointed out. The used methods are theoretically investigated and implemented in the existing tool-chain. Beside the possibility to obtain plans of high quality, the extended tools enable the user on the one hand to estimate the performance of a given (agile) Earth observation satellite respectively its capability to perform different observation schedules and on the other hand to measure the influence of specified satellite parameter on the resulting observation schedule during the design process.

Zugeordnete Arbeitspakete

AP 2300: Agile Erdbeobachtung
AP 2310: Konzepterstellung zur automatischen Erstellung von Beobachtungsplänen für agile Satelliten unter lagedynamischen Randbedingungen und Vermeidung von unerwünschten Aktuatorzuständen (CMG-Singularitäten)
AP 2320: Mathematische Modellbildung des Satelliten, insbesondere Lagedynamik, Manöver und Aktuatordynamik
AP 2330: Simulation in Matlab