Romie: A Domain-Independent Tool for Computer-Aided Robust Operations Management

Michael Saint-Guillain[a], Jonas Gibaszek[a], Tiago Vaquero[b], Steve Chien[b]

[a] *Université catholique de Louvain, Place de l'Universite 1, Ottignies-Louvain-la-Neuve, 1348, Belgium*
[b] *Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr, Pasadena, 91109, CA, USA*

Abstract

Romie is a decision support tool based on AI's latest advances in the domain of robust scheduling. Unlike all existing systems, the tool allows to (i) *visually model* the operational problem and context entirely (ii) optimize to find near-optimal schedules while *taking uncertainty into account* and (iii) deal with a combination of various *key performance indicators* (KPIs). It comes with a web user interface. Part or all of the modelled activities may be associated to random variables describing their stochastic durations, in order to produce schedules that are *robust* w.r.t. temporal uncertainty. Hence, depending on the pursued KPIs, the schedules maximize a combination of the following terms: the probability of satisfying the problem constraints, the expected return/efficiency, the expected outcome quality, and even the *operators' wellness* by minimizing its expected extra-hours. Initially developed for spatial exploration and demonstration in the context of Mars analog missions (i.e. missions on Earth that simulate condition and aspects of Mars missions), this *versatile* tool is here applied to operations management in both biotechnology manufacturing and robots parametrization in a cave exploration context.

1. Introduction

Project management realizes about 30% of the world gross product (Turner et al., 2010). However, most of existing studies have solely been done in machine scheduling environments (Herroelen and Leus, 2005). How to deal with processing time uncertainty when facing a larger, complex, scheduling problem which possibly involves multiple human and/or operators, unknown probability distributions, hard deadlines and exotic constraints?

In this paper, we introduce a visual tool, Romie, for computer-aided operations scheduling under uncertainty, and show how it has been successfully applied to three very different case studies of real world human operations management: 1) the *UCL to Mars 2018*[1] analogue mission that took place at the Mars Research Desert Station (Utah), 2) the modelling and scheduling of a manufacturing project in an real Belgian biotech company, and finally 3) the complex mission operations of the Jet Propulsion Laboratory (JPL) team in the DARPA Subterranean Challenge. The results obtained from our three case studies convey three very important messages: **(a)** Even for very complicated and various different operational contexts, a common modelling framework exists, being user friendly, visual, and rigorous at the same time; **(b)** Even for real sized problems, computer-optimized solutions outperform the schedules hand-crafted by field experts in general, and serve as a strong basis for decision making, as the deciders can always adapt and reuse these depending on external factors; **(c)** Schedules obtained while taking uncertainty into account systematically outperform that obtained from deterministic assumptions in terms of reliability and expected KPIs, while preserving most of the solutions quality; the latter result remains valid even when provided very bad representation of the uncertainty.

1.1. Description of the tool

Romie is an *advanced planning and scheduling* (APS) tool, i.e. a software system aiming at supporting the decision makers in their operations management and task scheduling, using a detailed domain model describing the operational context. Romie uses combinatorial optimization in order to generate and optimize robust plans for daily operations. The current key functionalities are depicted in Fig. 1:

- User friendly, *visual modelling* of the problem at stake, in its own operational context: human and physical resources, operational constraints, key performance indicators (KPIs), execution uncertainties.

- *Robust scheduling*: the optimization engine takes the time uncertainty on each task's duration into consideration, using modified-PERT distributions, yielding schedules with high probability of success.

[1]UCLouvain (Belgium) university. https://marsuclouvain.be

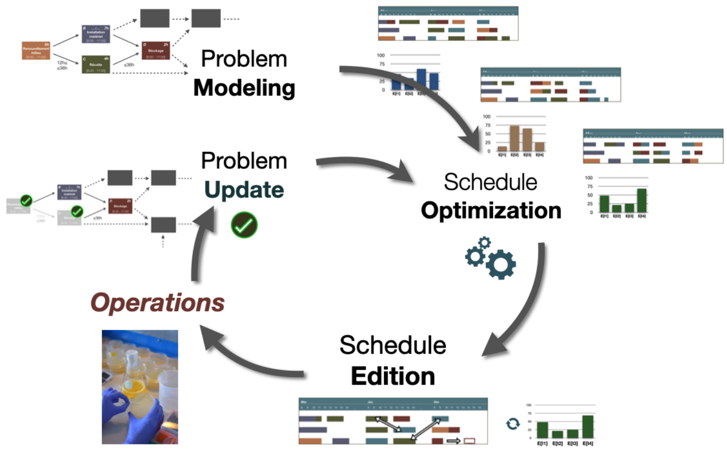

Figure 1: Principal functionalities of *Romie* tool.

- *KPI-guided scheduling*: The schedules are optimized while pursuing (a combination of) *various KPIs*, including success probability, expected cost, expected quality, and even operators wellness.

In a research domain in constant evolution, *Romie* integrates state-of-the-art advances in robust scheduling under uncertainty (Saint-Guillain et al., 2021). Future versions will enable online monitoring of the operations, keeping the schedule and the underlying model consistent with the current state of the system, allowing the user to adapt and reoptimize future decisions based on past outcomes.

Optimization engine. A local search (LS) based approach, exploiting well-known sequence neighborhood operators (relocate, 2-opt, swap, ...) and a simulated annealing meta-heuristic, is used to explore the solution space. The LS algorithm uses sample average approximation (Kleywegt et al., 2002) to evaluate the expected quality of a solution under time uncertainty. Each new solution is compared with the incumbent one in terms of its relative gain in each element of the ordered set of KPIs.

1.2. Timeline of Case Studies

This study presents a novel scheduling tool through three different case studies. Each case study happened sequentially, following and further validating different stages of the tool's development. Section 3 describes the *UCL to Mars 2018* mission case study, which marked the very first stage

Figure 2: Left: the Mars Desert Research Station in Utah. Right: extra-vehicular field operations.

of the technology, assessing its ability to deal with complex projects made of simple operational tasks, constraints and resources. Section 4 describes a case study which took place in 2019 and 2020, in collaboration with a Belgian biotech company. It permitted to extend our scheduling formalism and technology and reach a higher stage of applicability, in the complex real world industrial context. At the time these first two studies were conducted, the tool only consisted in a theoretical background, a modelling formalism and a versatile scheduling engine. Namely, there was no user interface (UI). The technology was validated, but not the ability of the end user to control and use it. Section 5 describes the third case study, in which human operators and robots from JPL collaborate in the final circuit of the DARPA Subterranean Challenge, scheduled for September 2021. *Romie* has been provided a brand new UI prototype, allowing (for the first time) an end-user to model the problem using a user-friendly visual interface, run optimization processes and visualize optimized schedules.

2. Related frameworks and software systems

There are many scheduling tools on the market. In a recent study, Abramov et al. (2016) describe some of the most common ones (MS Project, Primavera, Artemis, *etc*). However, one should pay attention to a fundamental difference between software systems under the very large denomination of ”*scheduling (or planning) tools*”, and the so-called ”*advanced planning and scheduling*” (APS) tools (Stadtler et al., 2015). When called APS, in addition to the ability of visualizing and manipulating schedules, the tool should come with an automated scheduling and optimization engine, able to generate solutions (*i.e.* schedules) based on a set of pre-specified operations and

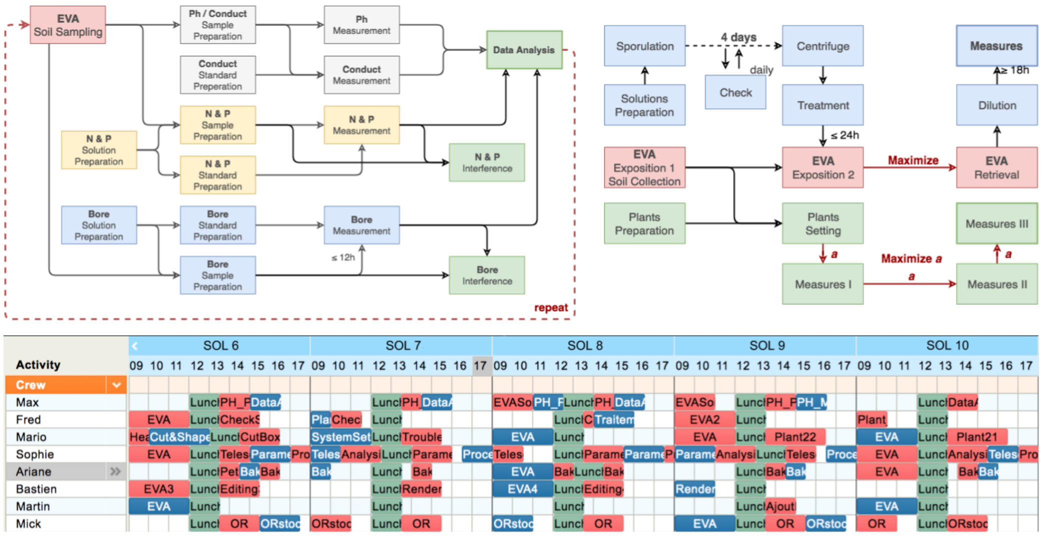

Figure 3: Top: visual models of two (out of seven) experiments conducted during the *UCL to Mars 2018* analog mission at MDRS. Top left: a soil analysis project in chemistry. Top right: a project mixing biology and botanic. Bottom: part of an optimized schedule followed the analog mission.

constraints (usually hidden to the user!). On the contrary, the vast majority of non-APS tools, such as MS Project, are mostly complex Gantt chart manipulation software systems. In other words, the user is still left with the initial problem of providing a scheduling solution, that is, the ordered sequences of tasks, each being assigned resources and scheduled times. In what follows, we focus on existing systems that provide both the modeling and the optimization capabilities.

Existing work either falls into *a)* being specifically designed for a particular application/mission or operational context or *b)* not having a generic, integrated optimization system to generate robust schedules (from a probabilistic point of view). Compared to the existing approaches, and to our knowledge, *Romie* provides the following *technological innovations*:

1. Domain-independent graphical modelling. Through a user-friendly interface, Romie provides to the user the ability of graphically drawing the structure and constraints of its planning and scheduling problem, including stochastic models for task durations.
2. Optimization under uncertainty. The optimization engine provided by *Romie* allows the end-user to rapidly generate schedules that are robust.

2.1. Planning and Scheduling in Space

The first planning and scheduling tools for space missions were dedicated software systems, specialized to specific application domains.

Johnston and Miller (1994) described the SPIKE system, a general framework for scheduling, developed by the Space Telescope Science Institute for NASA's Hubble Space Telescope. Other examples of aerospace scheduling tools and applications are: Chien et al. (1999), developed for scheduling the operations of a particular shuttle science payload (DATA-CHASER) with primary focus on solar observation; Jónsson et al. (2000) for the Deep Space One mission; Ai-Chang et al. (2004) for the Mars Exploration Rover mission; Chien et al. (2005) for NASA's Earth Observing One Spacecraft; and Cesta et al. (2007) for the Mars-Express mission. Chien (2012) provides a detailed survey on (semi-)automated planning & scheduling systems developed for space applications.

As the need for more generic approaches to support multiple mission and multiple domains increased, a planning/scheduling C++ library has been proposed: ASPEN (Fukunaga et al., 1997, Rabideau et al., 1999, Chien et al., 2000). At that time, ASPEN provided the elements that were commonly found in existing complex planning and scheduling systems, for example for generating operation schedules for the Rosetta orbiter Chien et al. (2021). In 2009, ESA's Advanced Planning and Scheduling Initiative (APSI) aimed at developing a general software framework for supporting development of AI planning and scheduling prototypes, for various types of space missions. The APSI is described in Steel et al. (2009).

Presented in Yelamanchili et al. (2020), the *Copilot* system for Mars 2020 Rover mission does have a modelling system called COCPIT, and a planner, but it is specifically designed for that mission.

2.2. Human Self-Scheduling in Space

Past space missions have had very limited experience in human self-scheduling. In fact, Marquez et al. (2019) states that current human operations, including extravehicular activities (EVAs), are *"carefully choreographed, and rehearsed events, planned to the minute by a large team of EVA engineers, and guided continuously from Earth"* (Bell and Coan, 2012, Miller et al., 2015). Whereas the delay in communications from Earth is negligible when activities remain near from Earth (*e.g.* from the ISS), sending (receiving) data to (from) Mars requires between 3 and 21 minutes. However, human operations on Mars are expected to be carried at a faster rate than current rover missions (Mishkin et al., 2007), which implies new planning strategies and tools that account for latency-impacted interactions (Eppler et al., 2013). In addition, future planetary EVAs are likely to be driven by science (Drake et al., 2010, Drake and Watts Kevin, 2014), requiring flexible adaptations according to scientific samples. In such context, future human space missions will have to enable some degree of crew autonomy and self-scheduling capabilities.

In Deans et al. (2017), a suite of software tools called Minerva is proposed in order to support operations planning and execution. Minerva and its components have been tested during several planetary and space simulation missions, including the BASALT research program (described in Brady et al., 2019) and four analog missions at NEEMO (Chappell et al., 2017, Marquez et al., 2017). Compared to Minerva, the key differences of our proposed tool *Romie*, in terms of functionalities, rely on the modelling interface and the scheduling optimization engine, which enable strategical a priori planning. In addition, the optimization is conducted while taking uncertainty into account. The Minerva suite is rather focused on tactical planning, including geospatial planning, which allows crew path planning and coordination using satellite maps. The strategical planning is assumed to be performed before the start of the mission, and is therefore not covered by the Minerva suite. However, even when a predefined schedule is provided prior to the start of the operations, it is very likely that the schedule will require online modifications as the operations go. Marquez et al. (2021) showed the limits of human self-scheduling when operators must solve and adapt the planning manually while taking hard constraints into account (not even thinking about uncertainty). By providing both a way to adapt the model and solve it using an embedded optimization engine, *Romie* can be seen as complementary to Minerva.

2.3. Human (Self-)Scheduling in the Industry

Before the ASPEN system described in Sec. 2.1, the idea of a generic (*i.e.* domain-independent) planning and scheduling library originally comes from the OZONE system of Smith et al. (1996), which has been applied to manufacturing, transportation and logistics.

Naturally, using such libraries requires specific programming skills. Such frameworks are therefore not a solution for enabling self-scheduling. Several domain-specific applications have been proposed using systems like OZONE (*e.g.* Smith and Lassila, 1993), but these generally appear to the user as a black box, in which the modification of any structural aspect of the scheduling problem at stake remains, if it is not impossible, a very complex task.

In Papavasileiou et al. (2007) and Petrides et al. (2014), the authors argue for the important role process simulation and scheduling tools in biopharmaceutical production, and in particular, the ability of performing what-if and/or sensitivity analysis in addition to optimization. However, because existing tools do not allow the end-user to model the production problem at stake, existing APS tools fail at meeting the resource and constraint structure involved in problems as complex as biomanufacturing. As a consequence, the stakeholders (when they can afford it) use expensive software systems, developed and bought specifically for their production processes (which hence cannot evolve in time without further expenses).

In this paper, we present an innovative general tool, which can be configured to meet a large range of operational contexts, from space missions to biomanufacturing, without domain-specific developments.

2.4. Romie

Recall the two technological innovations of Romie, presented in the beginning of this section: **a)** domain-independent graphical modelling and **b)** optimization under uncertainty. Unlike all existing tools[2], both modelling and modifying the problem is now made accessible to the end-user, which is critical for a reliable self-scheduling. Although being a hot research domain, *Romie* is the first APS tool to propose an integrated robust (*i.e.* under uncertainty) optimization engine. Having more robust (*i.e.* reliable) schedules, the end users are more likely to avoid last minute rescheduling. *Eventually, what-if analysis, as well as sensitivity analysis, become less relevant*: the solutions are optimized following directly the KPIs *expected values* and considering the uncertanties related to task execution.

We believe that both **a)** and **b)** provide significantly more autonomy to the end users, whom remain otherwise highly dependent of planning and scheduling experts. Whereas the empirical contribution of point *b* is assessed throughout this paper, the ability of the non-experts end-users to actually "self-schedule" using *a* remains to be empirically tested. This is left for

[2]Up to our knowledge, the MapGen tool presented in Ai-Chang et al. (2004) was one of the very first tools to propose a visual constraints editor. However, the latter was not generic, but specific to its application case, the NASA's MER mission.

further research on the field[3].

3. Robust Operations Management on Mars

The development of the Romie tool started with the *UCL to Mars 2018* project. Unlike most scheduling problems, operations in a space mission must be planned days ahead. Complex decision chains and communication delays prevent schedules from being arbitrarily modified, hence online reoptimization approaches are usually not appropriate. The problem of scheduling a set of operations in a constrained context such as the *Mars Desert Research Station* (MDRS, Fig. 2) is not trivial, even in its classical deterministic version. It should be seen as a generalization of the well-known NP-complete *job-shop scheduling problem* Lenstra and Kan (1979), which has the reputation of being one of the most computationally demanding (Applegate and Cook, 1991). Hall and Magazine (1994) reinforces the importance of mission planning, as 25% of the budget of a space mission may be spent in making these decisions beforehand, citing the Voyager 2 space probe for which the development of the a priori schedule, involving around 175 experiments, required 30 people during six months. Nowadays, hardware and techniques have evolved. It is likely that a super-equipped (*i.e.* with a brand new laptop) human brain suffices in that specific case. Yet, the problems and requirements have evolved too. Instead of the single machine Voyager 2, space missions have to deal with teams of astronauts. In fact, scheduling the activities in the ISS takes weeks, even for a team of experienced planning experts (Dempsey, 2017).

3.1. Scheduling a Space Mission under Uncertainty

The purpose of the *UCL to Mars 2018* analog mission being to simulate intensive scientific activities in a extra-planetary context, the mission was organized based on 7 different research projects to be conducted at the MDRS, from various fields including biology, particle physics, medicine, engineering, botanic, chemistry and finally AI. In total, more than 230 tasks were involved by the seven research projects, with at least as many constraints. In fact, the modelling each research project merged within a global problem was inevitable, since all the activities at MDRS depend on the same limited

[3] At the time of writing, a Mars analog mission (M.A.R.S. UCLouvain, scheduled April 2022) is preparing at the Mars Desert Research Station, Utah, during which *Romie* will be continuously used by analog astronauts during a simulated mission on the red planet.

and shared set of resources. Some projects required *extra vehicular activities* (EVAs), which for security reasons require at least three participants amongst the operators. EVAs usually take half a day, should be planned and approved days ahead and happen at most once a day. All operators then had their schedule linked to each others, even concerning research projects that do not require EVAs.

Fig. 3 shows the modelling of two such research projects, together with an excerpt of optimized schedule. At the MDRS however, computing an optimal schedule becomes significantly less attractive as problem data, such as the manipulation time of experiments, are different from their predicted nominal values. This is what we refer to as *uncertainty*. In a constrained environment with shared resources and devices, when they arise such temporal deviations can propagate to the remaining operations, eventually leading to global infeasibility, that is, a mission failure. Given a mission schedule, a central question is then the following: considering temporal uncertainty, what is the actual probability of success of the mission? In Saint-Guillain (2019), we investigated based on the real case study of a Mars analog mission, the impact of stochastic robust modelling against a classical deterministic approach on the reliability of a priori mission planning.

Uncertainty and Performance indicators. The main purpose of Saint-Guillain (2019) was to compare both deterministic and robust stochastic approaches to the problem of scheduling a set of scientific tasks under processing time uncertainty, in the operational context of a Martian planetary habitat. We empirically showed that taking uncertainty into account while optimizing the schedules allows significant gains on average when applied on real instances involving the constraints faced and objectives pursued during a two-week Mars analog mission. The objectives were both optimizing the mission's success probability, namely the *robustness*, in terms of meeting the operational constraints and deadlines, and maximizing the total scientific outcome, as a linear combination of specific metrics, designed according to the scientific objectives of each of the seven research projects. Therefore, the optimization engine was computing a near-optimal schedule when optimizing the robustness KPI first, and expected scientific outcome second.

Uncertainty on the Uncertainty. Contrary to other application domains which may come with a huge number of observations, projects in human operations management are hardly repeatable. However, accurate estimations of the

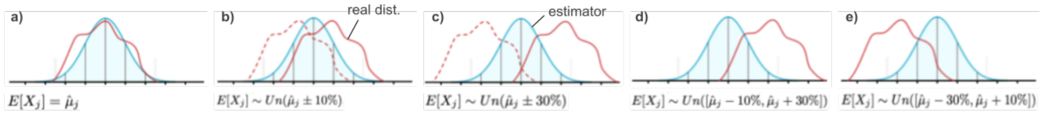

Figure 4: Varying the quality of the probability distributions, leading to five different experimental assumptions. Blue: estimator distribution, used at optimization stage. Red, both plain and dotted: real hidden distributions, revealed at execution stage.

probability distributions require a significant amount of observations, which is often impossible in practice. One of the main contributions of our approach is that it computes decisions while taking uncertainty into account, in terms of activity durations. In the context of human scheduling, these decisions must then be taken solely based on the expert's estimation of that uncertainty, that is, on the stakeholder's field experience, instead of statistics.

For that reason, the computational experiments did also take into account uncertainty on the stochastic knowledge itself, by considering the real distributions as unknown (or hidden). In Saint-Guillain (2019) the same computational experiments were reiterated under five different assumptions about the estimator's quality. The concept is illustrated in Fig. 4: five hypothetical couples of both estimator and real distributions are drawn. The estimator distribution is simply the distribution used to describe the duration of an activity. It represents the current knowledge one has about the activity's uncertainty. The estimator distributions are the only information available at optimization phase. Real (hidden) distributions are only revealed when the computed schedules, therefore optimized considering blue distributions, were executed in a simulation. The shapes of the real distributions were always randomly generated, while controlling the resulting mean. In the first case **a)**, the estimators (which used to be normal distributions) were of very good quality, since their means always coincide with that of the real distributions (even if the shapes differ). In the third case **c)**, these were of bad quality, since each estimator mean may be up to 30% away from that of the associated real distribution. Finally, the fourth **d)** (*resp.* fifth) case stands for the situation in which all durations are underestimated (*resp.* overestimated, **e**) in general. The results revealed that, even when using very bad approximations of the probability distributions, the computed solutions significantly outperform those obtained from a classical deterministic formulation, while preserving most of the solution's quality.

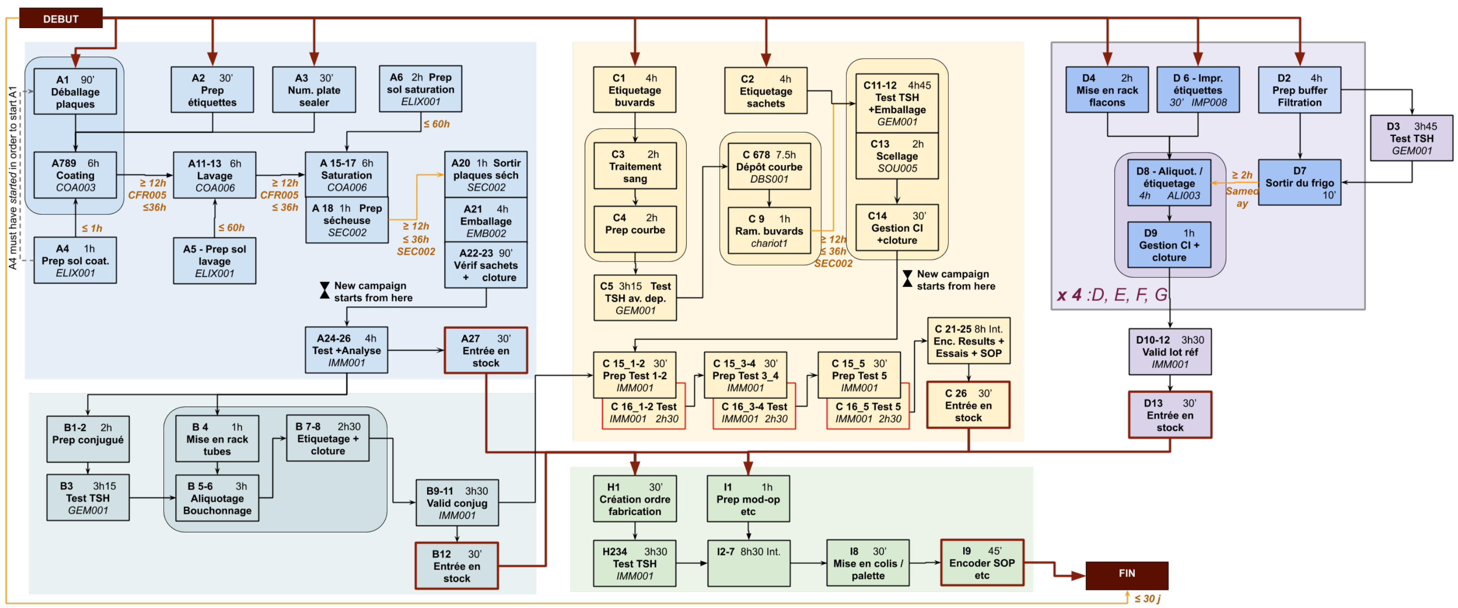

Figure 5: A visual model that fully describes the tasks and constraints involved in one production campaign of the biotech company's most popular product.

3.2. Experiments and Results

The aforementioned paper gives extensive details on both the experimental plan and results. In a nutshell, we observed that as the accuracy of the probability distributions that describe the project task durations varies from really accurate (no under/over estimation on average) to very bad (30% under/over estimations on average), the proportion of the simulations in which the schedules computed based on a classical deterministic model lead to a successful execution varies from 5% to 8% only. Using our probabilistic model, these success rates increase significantly, between 70% and 90%. Two additional assumptions were tested, in which the durations were either systematically underestimated, or systematically overestimated. The first case is naturally catastrophic, leading to success of approximately 0.1% for the deterministic schedules, and 22% thanks to our stochastic model. The second case is really interesting, as it describes a very common behaviour of managers who have to face time uncertainty, which consists in systematically considering a duration larger than what they believe the tasks is likely to last. In this particular context, our simulations showed that our stochastic model increases the schedules' robustness from 34% to 95%.

Benefits and Price of Robustness. Whereas we principally focused so far on the robustness, in practice one is also necessarily interested in the *outcome* or, alternatively, the price (or cost) of the a priori solution. For example, suppose a solution A with a nice 90% robustness, but costing 100$. On the other hand, you are provided an alternative solution B, obtained with a good, old fashion, deterministic model. Solution B appears much less reliable with only 30% robustness. However, B costs only 10$. Would you really go for a solution A, which is three times more reliable, but ten times more expensive? Probably, if you are planning a space mission or a surgery. Probably not, if you are planning financial investments.

In the case of our space mission, the cost, or outcome, actually matters and is expressed in terms of scientific goals and preferences to be optimized. For instance, the top left model of Figure 3 stipulates that in order to maximize the scientific outcome, the delays between some of the activities must be maximized. Remark that such an objective is in opposition with our robustness criterion, as one the main success issue relates to completing the mission in time.

In terms of robustness, the experimental results clearly indicate the benefits of our probabilistic model over a deterministic one, as even in the ideal case of all durations being overestimated on average, the resulting schedules reveal three times more reliable (succeed in 95% of the cases). Furthermore, independently of the estimators accuracy, the relative difference in terms of the scientific outcome KPIs, between schedules produces from both models, is of only 7% on average. In other words, even all durations being overestimated and in our MDRS context, *a deterministic schedule yields 7% more science on average, whereas it is three times more likely to fail the mission.* This is the value of perfect information over uncertainty. On the other hand, a success rate of 95% can be reached by sacrificing 7% of the outcome. This is the benefits, and price, of robustness.

4. Biotech Manufacturing

As the day humans will live on Mars is still far ahead, we wanted to extend our tool as well as the underlying technology to tackle significantly different operational context, others than human space missions. A Belgian company, specialized in biotechnology product manufacturing, accepted to collaborate on the concrete project of modelling the scheduling problem involved in the manufacturing of one of their most popular products, and eventually solving this scheduling problem, at different scales.

4.1. Modelling in the Industrial Real World Contexts

Fig. 5 shows how their production problem was modelled, using the exact same *visual formalism* than that used in the context of the *UCL to Mars 2018* mission. More precisely, the diagram depicted only represents one single production campaign, involving around 85 tasks. However, for efficiency reasons the company would usually run up to three production campaigns in parallel, whereas the operational human and physical resources remain fixed.

As the operational context in a biotech company significantly differs from that of a space station, various additional exotic constraints had to be added to the formalism, and the other scheduling optimization technology extended consequently. For instance, running several production campaigns in parallel involve specific constraints, stating how these may overlap or not. Naturally, we had to cope with complex operator worktime management, such as weekends, days off, part times, *etc.*. In particular, the concept of extra working time, whenever an operator must remain on site later than normally accepted, revealed of major importance. Finally, as the production process requires many tasks to be double checked, constraints of the form “task A cannot be executed by the same operator than task B” were mandatory too, and special constraints related to physical resources had to be designed as well. These are just examples amongst the large number of additions that extended the initial technology. Fig. 6 shows an example of solution for three production campaigns in parallel.

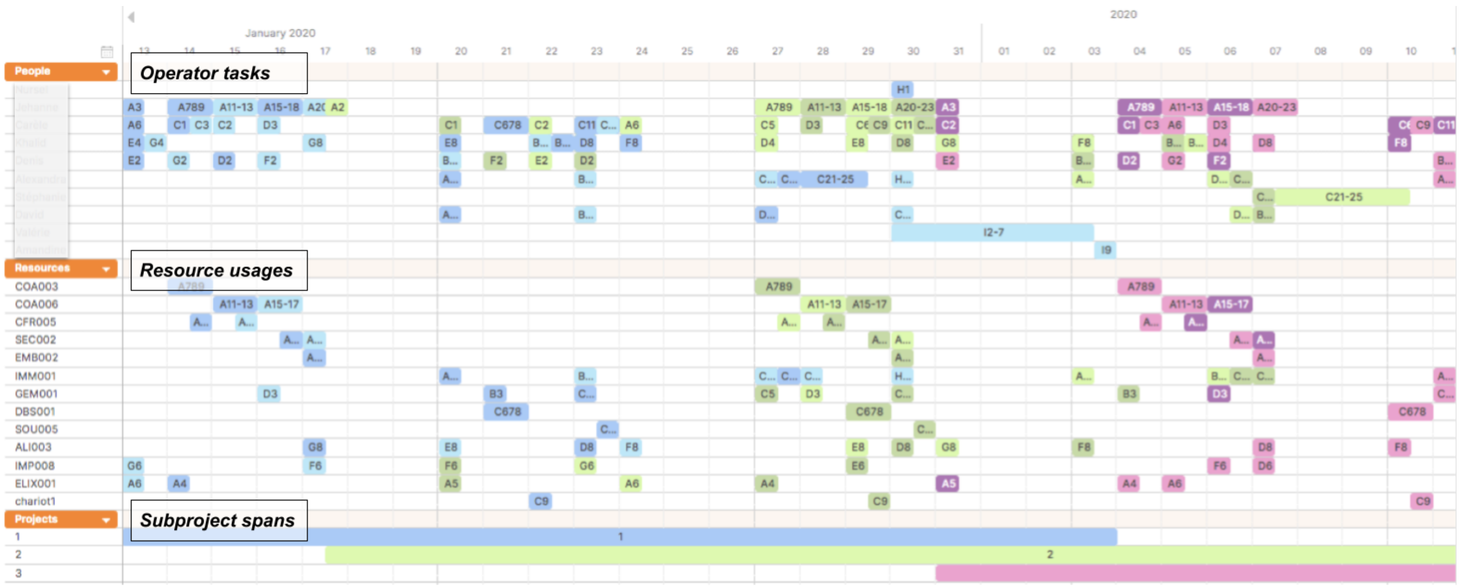

Figure 6: A possible optimized solution to the problem described by the model depicted in Fig. 5 when three production campaigns are scheduled in parallel. The visualization includes the operators' task sequences as well as the resource (machines) usages. We see how the campaigns overlap at the bottom with the blue, green and purple bars: each denotes a campaign.

4.2. Problem Modelling made Simple

The modelling formalism used to communicate with our collaborators from the biotech company, namely the language used to describe their scheduling problem, is based on simple diagrams such as depicted in Fig. 5. In practice, once the problem visually modelled, we translated the diagrams in a mathematical language accepted by our scheduling engine. It is worth noting that the time needed by our collaborators, being not scheduling experts, to master the proposed visual modelling formalism (and autonomously draw part of the models) revealed to be of five to ten one-hour meetings only.

4.3. Computing Robust production Schedules

Computational experiments involve from one to three production campaigns. For each context, defined by the amount of campaigns and whether the stochastic or the deterministic model was used, 10 solutions were optimized. Average results are listed in Table 1. An example of such computed schedule, for three campaigns in parallel, is shown in Fig. 6.

Optimizing Wellness: Stress Aversion. Cost-based KPIs, such as minimizing total production makespan, or quality-based KPIs such as the metric of total scientific outcome considered in the *UCL to Mars 2018* case study, are classical objectives to be pursued. In a human context however, wellness and stress aversion are key concepts that should be considered as important as raw efficiency in the middle and long term. Our collaborators from the biotech company recognized that a significant part of their employees' stress can be attributed to unexpected deviations, resulting in delays which, eventually, force the production team to do extra-hours in order to stick to the constraints and deadlines. Consequently, it has been decided to consider the *expected total number of extra-hours* as key performance indicator.

Performances of the deterministic model. The average reliability of the solutions optimized under deterministic model significantly fall as the complexity of the problem increases. Move from one production campaign to three campaigns, the success rate under the over-estimation context falls from 40% to 12.5%, when the extra-hours (EHs) KPI is minimized prior to the makespan. In fact, schedule with less planned EHs are more flexible, more likely to be able to absorb unexpected delays, and thus more reliable in the end. In particular, when minimizing EHs the computed schedules come with 0.0 planned EH whilst, eventually, around 4.3∼5.9 EHs are required on average, for one production campaign. Given three campaigns, significantly larger deviations are observed from the initially planned EHs of only 1 hour on average, which increases up to 11.3∼23.6 hours. Larger deviations of the planned versus observed makespan KPI are also measured as the size of the problem increases.

Performances of the stochastic model. Compared to the solutions obtained when all durations are considered as perfectly known in advance, namely when using a deterministic model, the average measured performances of the schedules computed in light of uncertainty are ridiculously obvious. Whereas the success rate can be maintained above 96% for three campaigns (instead of 2.5%!), the price of this robustness as measured by the average difference in the makespan cost KPIs, is of only 13% (38.1 to 43.2 days) when optimizing makespan first. When minimizing EHs first, makespan second, taking uncertainty into account at optimization stage leads to schedules having 98.7% chances of success on average, instead of 12.5%, with significantly less work stress as the measured extra-hours are of 3.8∼9.9 hours, instead of 11.3∼23.6 hours. The price of robustness, in terms of production efficiency, is however higher at it is now of +29% makespan (45.5 to 58.9 days). Yet, anyone would be surprised by a manager that decides to save the 29% and goes for a schedule having only 12.5% chances to succeed.

		% Success			Extra-Hours				Makespan (days)			
		Exact	Under	Over	Plan	Exact	Under	Over	Plan	Exact	Under	Over
1C Det.	Min. makespan	6.6	0.9	17.1	9.9	9.7	11.3	10.0	15.1	15.7	16.6	15.4
	Min. extra-hours	60.9	37.5	40.2	0.0	4.6	5.9	4.3	17.2	19.8	20.7	18.5
1C Stoch.	Min. makespan	98.4	83.7	98.0	1.4	4.1	6.1	3.1	17.2	17.2	17.3	17.2
	Min. extra-hours	99.9	99.6	100.0	1.0	1.9	2.7	1.0	26.9	27.1	27.4	27.0
3C Det.	Min. makespan	0.5	0.0	2.5	16.6	29.1	34.6	26.7	38.1	41.9	42.5	40.1
	Min. extra-hours	14.2	2.2	12.5	1.0	15.7	23.6	11.3	45.5	46.4	46.8	45.9
3C Stoch.	Min. makespan	96.2	61.1	96.0	6.2	13.0	20.0	10.3	43.2	43.2	43.3	43.2
	Min. extra-hours	99.4	97.7	98.7	3.0	6.4	9.9	3.8	58.9	58.9	58.9	58.9

Table 1: Biotech manufacturing case study, involving one (1C) and three (3C) production campaigns, optimizing under either deterministic (Det.) or stochastic model (Stoch.). We consider 3 different assumptions about the stochastic knowledge: exact mean, 10% under-estimations and 10% over-estimations, as described in Section 3: first, fourth and fifth in Fig. 4, this time with $E[X_j] \sim U(\mu_j - 3\%, \mu_j + 10\%)$ and $E[X_j] \sim U(\mu_j - 10\%, \mu_j + 3\%)$. Statistics include the percentages of simulations in which the optimized schedules (% Success) are found to respect all the problem constraints when executed under the "hidden" uncertainty (red in Fig. 4). The *Plan* columns indicate the KPI values as predicted by the *a priori* schedule, as opposed with the average KPI values observed during simulations (other columns). *Minimize makespan*: optimization done while minimizing the makespan first, extra-hours second. *Minimize extra-hours*: extra-hours first, makespan second.

5. DARPA Subterranean Challenge

NASA JPL's Team CoSTAR[4] is developing new technologies that are critical for enabling autonomous multi-robot exploration of large and unknown underground voids. One example of the application of these technologies is the DARPA Subterranean Challenge (SubT) where terrestrial cave exploration can be seen as an analogue exploration mission for planetary subsurfaces (e.g. Lunar and Martian caves), and as an application domain to prove grounds for future space technologies. In SubT, robot teams are required to rapidly map, navigate, and search underground environments including natural cave networks, tunnel systems, and urban underground infrastructure for particular objects of interest (*e.g.* mannequin survivors, backpacks, cell phones, helmets, *etc.*), called *artifacts*. Subterranean environments pose significant challenges for manned and unmanned operations due to limited communication and situational awareness. In the SubT competition in particular, only a single human operator is allowed to interact with the robotic team. CoSTAR's robotic team consists of more than four robots with wheeled, legged, and flying mobility.

Operating multiple robots with different capabilities in kilometer-long underground environments can go beyond the cognitive capacity of a single human supervisor – even with advanced autonomy in place. SubT operations may involve cognitively demanding tasks such as monitoring 3D mapping of the environment and localization accuracy, establishing communication links between robots, assessing location and health of all robots, and submitting detected artifacts within the allotted competition time. In order to facilitate operations during the SubT competition, CoSTAR team has developed the Copilot MIKE (Kaufmann et al., 2021), an autonomous assistant for human-in-the-loop multi-robot operations. During complex and potentially stressful exploration missions, MIKE helps by planning operation tasks related to setting up and commanding the robots, while maintaining a bearable workload and high situational awareness.

In this experiment we study the use of *Romie* to i) model the tasks and constraints that are required during setup time and during competition, and to ii) support the operator by scheduling the tasks in a way that maximizes the robustness. This case study has principally differs from the two previous ones by the scale of its time horizon to manage. Whereas the schedules at

[4]DARPA Subterranean Challenge Team CoSTAR. https://costar.jpl.nasa.gov

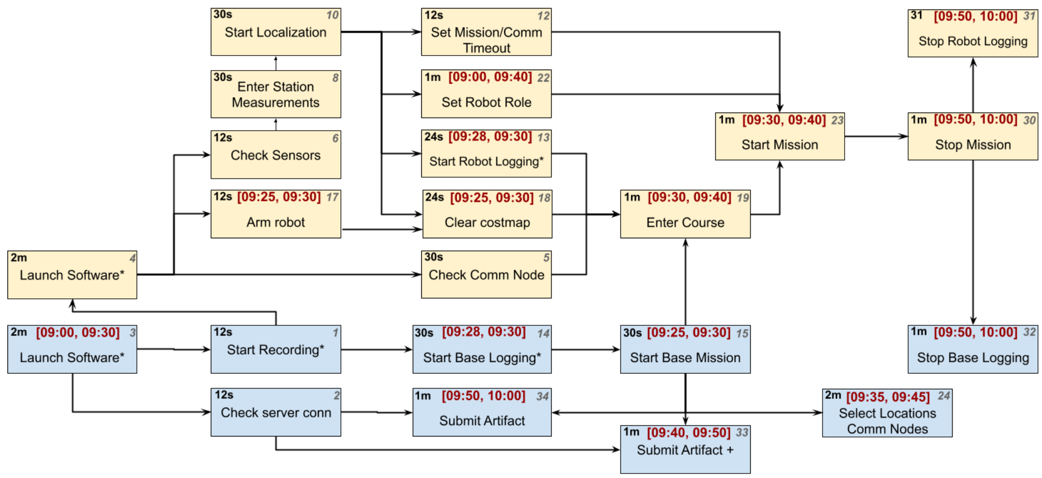

Figure 7: The DARPA SubT problem structure, as designed by hand using a classical diagram editor.

MDRS, as well as in the context of biotech production, typically involve several weeks, the DARPA SubT lasts only one hour. In particular, the problem at stake consists here at both 1) setting and getting all the exploration robots for starting the mission (setup time) and 2) deploying the team of robots to explore the target underground environment. The scheduling problem involves setup tasks having nominal durations ranging between 10 and 120 seconds.

5.1. Problem Modelling

Figure 7 shows a graphical drawing of the problem structure. This diagram, as well as those shown in Figures 3 and 5, have been drawn using *GoogleDocs Drawings*. The model involves two principal groups of activities: the *Base group* (in blue) and the *Robot group* (in yellow). The Base group includes unique activities that are common to all robots. The Robot group represents all the activities specific to one robot. If we have 7 robots, there are 7 duplicates of each activity from the Robot group. The main challenge results in the fact that everything must be scheduled in order to fulfill the time constraints, depicted in red. For instance, the *Start Robot Logging* activity must append between 9:28am and 9:30am, for every robot. Finally, the human operator is assumed to be able to carry on up to 4 activities at the same time, whereas some activities (denoted by * in Fig. 7) can be entrusted to MIKE, which could follow up to 5 activities in parallel. It is important to note that the only key performance index (KPI) here is the probability of success, that is, the probability that the schedule actually respects all the temporal constraints shown in red in Fig. 7. In fact, contrary to previous case studies, the problem at stake here is no longer an optimization problem but a constraint satisfaction problem, in the sense that there is no other goal to be pursued than satisfying these constraints.

The Modelling User Interface. Our tool now integrates a visual modelling user interface, depicted in Figure 8. The UI allows to describe the scheduling problem structure, in terms of temporal and operational constraints. It also allows to input constraints and activity properties, which were not present in the hand-drawn diagram of Fig. 7. In particular, the model allows for modified-PERT distributions, a probability distribution widely used in risk analysis (Kamburowski, 1997), which has the advantage of enabling asymmetric bounded probability distributions.

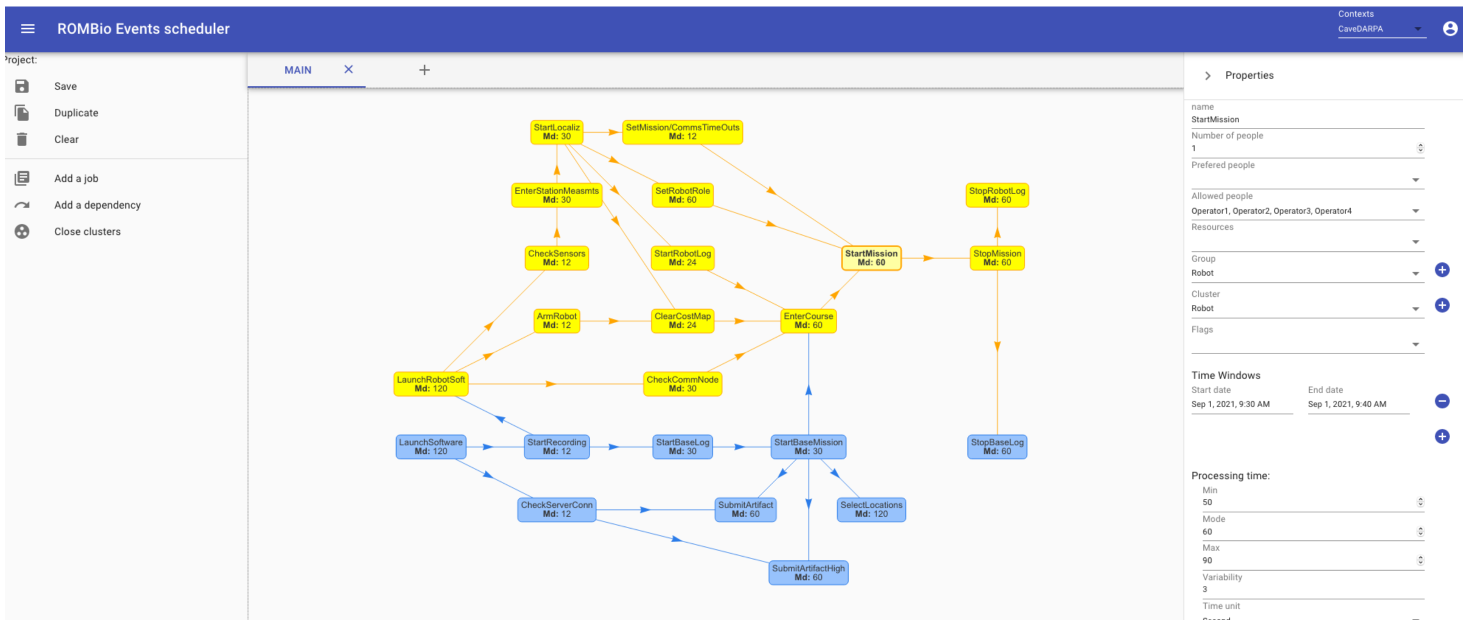

Figure 8: Modelling using the *Romie* tool's user interface. Amongst the different parameters that define an activity: allowed operators, used resources, execution time windows, parameters defining the duration's probability distribution (PERT).

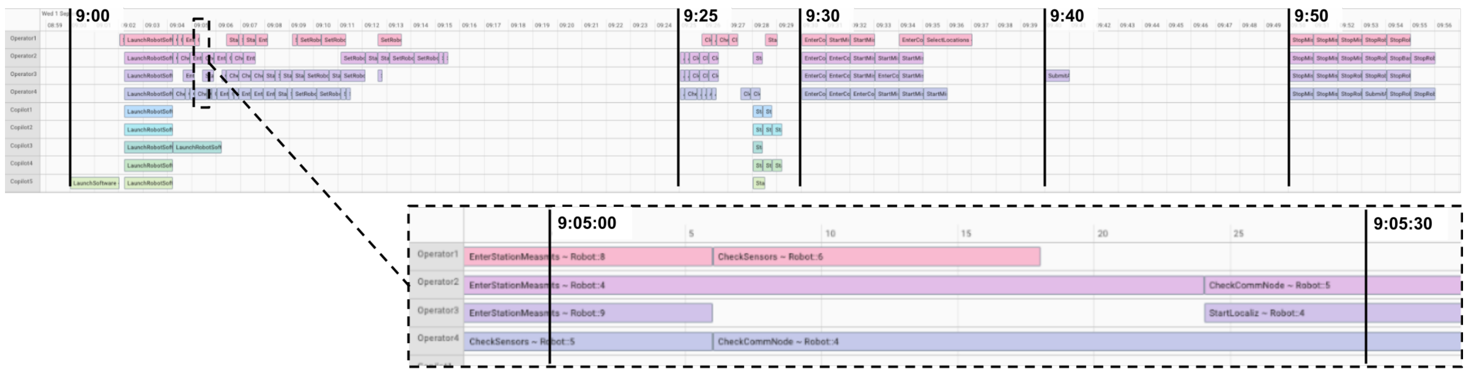

Figure 9: An example schedule solving the DARPA Subterranean robot problem, showed in the *Romie* tool's schedule visualization interface. Top: 1 robot problem. Middle: 10 robots problem. Each line corresponds to an operator timeline (4 humans, 5 copilots). The interface allows the user to zoom at will, has showed in the bottom of the figure.

5.2. Computing Robust Schedules

Our scheduling engine exploits parallelization paradigm to achieve reasonable computation times *Romie* is capable of exploiting parallel computing. On a 64 cores computation cluster (composed of two AMD EPYC 7302), finding feasible solutions (*i.e.* schedules that fulfil the temporal constraints listed in Fig. 7) takes less than 15 seconds only for problems with up to 15 robots. Up to 18 robots, it takes less than a minute to find its first feasible solution. Once an initial solution has been found to be deterministic feasible, meaning that it is feasible when considering all its durations as deterministic, the engines switch on the probabilistic optimization mode and pursue its duty while optimizing with respect to the expected KPIs: here the probability of success only. No deterministic feasible solution was found for 19 robots or more (in less than five minutes). An optimized schedule is depicted in Fig. 9, using the tool's integrated visualization interface.

Still exploiting 64 cores, for problems involving up to 12 robots *Romie* finds *optimal* solutions in less than 5 seconds. Of course, the nature of the scheduling engine embedded in *Romie*, a local search based solution framework, naturally prevents from providing any optimality proof in general. For the DARPA SubT challenge however, the only optimized KPI being the success probability, solutions with 1.0 success probability are necessarily optimal. In our case, Romie finds such solutions for instances involving up to 15 robots, in which 20 seconds only are required to find a first deterministic feasible solution, and optimal solutions are found within approximately 2 minutes.

Average Performances of the Computed Schedules. As usual, the optimized robust schedules are compared to solutions optimized based on deterministic assumption, therefore measuring the average gain at optimizing based on a probabilistic model against a classical, deterministic one. The average results are provided in Table 2. Solutions optimized under the stochastic model significantly outperform that of the deterministic one, in terms of reliability (*i.e.* robustness), that is, the average percentage of simulation success under the three different assumptions made on the quality of the stochastic knowledge (*i.e.* accuracy of the chosen PERT distribution parameters for each activity duration).

Bots	1	4	6	10	15	16	17	18
Exa	91.0	82.8	67.8	31.9	0.1	0.0	0.0	0.0
Und	90.0	80.8	58.5	17.4	0.0	0.0	0.0	0.0
Ovr	93.3	88.1	81.4	52.7	9.6	7.0	4.8	0.2
Exa	100	100	100	100	100	95.6	61.5	0.3
Und	100	100	100	100	93.0	40.7	1.3	0.0
Ovr	100	100	100	100	100	100	97.9	56.4

Table 2: DARPA SubT challenge case study, involving up to 18 robots. We consider three different assumptions about the stochastic knowledge: exact mean (Exa), [-3%,+10%] under-estimations (Und) and [-10%,+3%] over-estimations (Ovr), as described for Table 2. First raws give the average percentage success, when schedules are optimized using a deterministic model. The second set of raws give average results when *Romie* uses the stochastic model.

6. Conclusions and Future work

In this paper we presented *Romie*, a state-of-the-art robust scheduling tool, based on the principle of optimizing under time uncertainty. We described three very different application cases that were handled with our tool, showing the versatile aspect of *Romie*, which allows to model and solve scheduling problem despite the different operational contexts. The more recent case study shows that end users are actually able to visually describe the scheduling problem at stake, and further visualize optimized solutions. Once again, the benefits of using a probabilistic modelling approach, taking the time uncertainty of activities durations into account at optimization stage, are clearly confirmed by the empirical average gains compared to a classical deterministic approach. Our approach also allows to significantly decrease extra hours and deviations from a priori decisions, hence reducing the operators' stress load in manufacturing context.

A new risk-aversion paradigm. Up to now, the classical paradigm in operations management to cope with uncertainty was based on "*what if analysis*" and "*sensitivity analysis*". What if analysis consists in optimizing a set of solutions, each under a predefined scenarios (*e.g.* best-case, average-case, worst-case). As showed in Saint-Guillain et al. (2021), from a theoretical point of view this approach is a fundamentally flawed, as it violates the natural nonanticipativity constraints and therefore, may *arbitrarily underestimate* the expected behavior of the solution (that is, the *risk*!). On the contrary, the optimization engine embedded in *Romie* is a completely different approach, proven to *never underestimate the risk*. Sensitivity analysis consists in assessing the average behavior of a solution when subject to stochastic variations (*e.g.* using Monte Carlo simulation). Given a particular schedule, a sensitivity analysis will therefore approximate the expected quality (*i.e.* expected KPIs) of the provided schedule under uncertainty. This however does not help at finding the right schedule in the first place! Because its optimization engine directly take uncertainty into account at schedule generation and optimization, the solutions computed by *Romie* directly optimize their response to a sensitivity analysis. In other words, *Romie*'s optimization engine turns both the (inadequate) *what if analysis* and the (incomplete) *sensitivity analysis*, deprecated.

Future researches and development directions. Our tool currently integrates the key functionalities for *a priori* robust scheduling under uncertainty. The next logical step will naturally be to integrate online management functionalities, namely *monitoring* and *reoptimization*. Online monitoring of the operations aims at updating the schedule, as well as the underlying model, in order to keep them consistent with the current state of the operations. It eventually allows the possibility for adapting and reoptimizing future decisions, in light of past decisions and outcomes. These additional features will be useful for the scheduling of future research and operational projects. Applications to biotech and pharmaceutical online manufacturing problems are already running at the time of writing, as well as the next circuit of the DARPA Subterranean Challenge (September 2021). Finally, a simulated mission in a Mars analog habitat is planned in the near future with the M.A.R.S. UCLouvain mission (April 2022), during which a team of 8 analog astronauts will be assessed on their ability of self-schedule their own mission (both a priori and on-the-fly) using *Romie*. In fact, past missions (*e.g.* UCL to Mars 2018, Saint-Guillain 2019) have shown the importance of online reoptimization and, in particular, the need for the crew to autonomously adapt their science projects to unforeseen events.

Acknowledgements. We thank the *UCL to Mars 2018* crew: Bastien Baix, Frédéric Peyrusson, Martin Roumain, Ariane Sablon, Mario Sundic, Sophie Wuyckens and Maximilien Richald. We thank our collaborators from the Belgian biotech company *Zentech*: Isabelle Sansdrap, Nursel Yilmaz. We thank the Mars Society, Shannon Rupert and Mission Support personnel, for creating the conditions for realistic Mars analog sojourns. The research was partially carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004).

References

Abramov, I., Poznakhirko, T., Sergeev, A., 2016. The analysis of the functionality of modern systems, methods and scheduling tools, in: MATEC Web of Conferences, EDP Sciences. p. 04063.

Ai-Chang, M., Bresina, J., Charest, L., Chase, A., Hsu, J.J., Jonsson, A., Kanefsky, B., Morris, P., Rajan, K., Yglesias, J., et al., 2004. Mapgen: mixed-initiative planning and scheduling for the mars exploration rover mission. IEEE Intelligent Systems 19, 8–12.

Applegate, D., Cook, W., 1991. A computational study of the job-shop scheduling problem. ORSA Journal on computing 3, 149–156.

Bell, E., Coan, D., 2012. A review of the approach to iss increment crew eva training, in: AIAA SPACE 2007 Conference & Exposition, p. 6236.

Brady, A.L., Kobs Nawotniak, S.E., Hughes, S.S., Payler, S.J., Stevens, A.H., Cockell, C.S., Elphic, R.C., Sehlke, A., Haberle, C.W., Slater, G.F., et al., 2019. Strategic planning insights for future science-driven extravehicular activity on mars. Astrobiology 19, 347–368.

Cesta, A., Cortellessa, G., Denis, M., Donati, A., Fratini, S., Oddi, A., Policella, N., Rabenau, E., Schulster, J., 2007. Mexar2: Ai solves mission planner problems. IEEE Intelligent Systems 22, 12–19.

Chappell, S.P., Beaton, K.H., Newton, C., Graff, T.G., Young, K.E., Coan, D., Abercromby, A.F., Gernhardt, M.L., 2017. Integration of an earth-based science team during human exploration of mars, in: 2017 IEEE Aerospace Conference, IEEE. pp. 1–11.

Chien, S., 2012. A generalized timeline representation, services, and interface for automating space mission operations, in: SpaceOps 2012, p. 1275459.

Chien, S., Rabideau, G., Knight, R., Sherwood, R., Engelhardt, B., Mutz, D., Estlin, T., Smith, B., Fisher, F., Barrett, T., et al., 2000. Aspen-automating space mission operations using automated planning and scheduling, in: SpaceOps 2000, AIAA Press.

Chien, S., Rabideau, G., Willis, J., Mann, T., 1999. Automating planning and scheduling of shuttle payload operations. Artificial Intelligence 114, 239–255.

Chien, S., Sherwood, R., Tran, D., Cichy, B., Rabideau, G., Castano, R., Davis, A., Mandl, D., Frye, S., Trout, B., et al., 2005. Using autonomy flight software to improve science return on earth observing one. Journal of Aerospace Computing, Information, and Communication 2, 196–216.

Chien, S.A., Rabideau, G., Tran, D.Q., Troesch, M., Nespoli, F., Perez-Ayucar, M., Costa-Sitja, M., Vallat, C., Geiger, B., Vallejo, F., Andres, R., Altobelli, N., Kueppers, M., 2021. Activity-based scheduling of science campaigns for the rosetta orbiter. Journal of Aerospace Information Systems 18:10, 711–727. URL: `https://doi.org/10.2514/1.I010899`.

Deans, M., Marquez, J.J., Cohen, T., Miller, M.J., Deliz, I., Hillenius, S., Hoffman, J., Lee, Y.J., Lees, D., Norheim, J., et al., 2017. Minerva: user-centered science operations software capability for future human exploration, in: 2017 IEEE Aerospace Conference, IEEE. pp. 1–13.

Dempsey, R.C., 2017. The International Space Station: Operating an Outpost in the New Frontier. Government Printing Office.

Drake, B.G., Hoffman, S.J., Beaty, D.W., 2010. Human exploration of mars, design reference architecture 5.0, in: 2010 IEEE Aerospace Conference, IEEE. pp. 1–24.

Drake, B.G., Watts Kevin, D., 2014. Human exploration of mars design reference architecture 5.0, addendum# 2 .

Eppler, D., Adams, B., Archer, D., Baiden, G., Brown, A., Carey, W., Cohen, B., Condit, C., Evans, C., Fortezzo, C., et al., 2013. Desert research and technology studies (drats) 2010 science operations: operational approaches and lessons learned for managing science during human planetary surface missions. Acta Astronautica 90, 224–241.

Fukunaga, A., Rabideau, G., Chien, S., Yan, D., 1997. Towards an application framework for automated planning and scheduling, in: 1997 IEEE Aerospace Conference, IEEE. pp. 375–386.

Hall, N.G., Magazine, M.J., 1994. Maximizing the value of a space mission. European journal of operational research 78, 224–241.

Herroelen, W., Leus, R., 2005. Project scheduling under uncertainty: Survey and research potentials. European journal of operational research 165, 289–306.

Johnston, M., Miller, G., 1994. Intelligent scheduling. ch. Spike: Intelligent scheduling of Hubble Space Telescope observations , 391–422.

Jónsson, A.K., Morris, P.H., Muscettola, N., Rajan, K., Smith, B.D., 2000. Planning in interplanetary space: Theory and practice., in: AIPS, pp. 177–186.

Kamburowski, J., 1997. New validations of pert times. Omega 25, 323–328.

Kaufmann, M., Vaquero, T., Correa, G., Otsu, K., Ginting, M., Beltrame, G., Agha-Mohammadi, A.A., 2021. Copilot mike: An autonomous assistant for multi-robot operations in cave exploration, in: IEEE Aerospace Conference.

Kleywegt, A.J., Shapiro, A., Homem-de Mello, T., 2002. The sample average approximation method for stochastic discrete optimization. SIAM Journal on Optimization 12, 479–502.

Lenstra, J.K., Kan, A.R., 1979. Computational complexity of discrete optimization problems, in: Annals of discrete mathematics. Elsevier. volume 4, pp. 121–140.

Marquez, J.J., Edwards, T., Karasinski, J.A., Lee, C.N., Shyr, M.C., Miller, C.L., Brandt, S.L., 2021. Human performance of novice schedulers for complex spaceflight operations timelines. Human Factors , 00187208211058913.

Marquez, J.J., Hillenius, S., Kanefsky, B., Zheng, J., Deliz, I., Reagan, M., 2017. Increasing crew autonomy for long duration exploration missions: self-scheduling, in: 2017 IEEE Aerospace Conference, IEEE. pp. 1–10.

Marquez, J.J., Miller, M.J., Cohen, T., Deliz, I., Lees, D.S., Zheng, J., Lee, Y.J., Kanefsky, B., Norheim, J., Deans, M., et al., 2019. Future needs for science-driven geospatial and temporal extravehicular activity planning and execution. Astrobiology 19, 440–461.

Miller, M.J., McGuire, K.M., Feigh, K.M., 2015. Information flow model of human extravehicular activity operations, in: 2015 IEEE Aerospace Conference, IEEE. pp. 1–15.

Mishkin, A., Lee, Y., Korth, D., LeBlanc, T., 2007. Human-robotic missions to the moon and mars: operations design implications, in: 2007 IEEE Aerospace Conference, IEEE. pp. 1–10.

Papavasileiou, V., Koulouris, A., Siletti, C., Petrides, D., 2007. Optimize manufacturing of pharmaceutical products with process simulation and production scheduling tools. Chemical Engineering Research and Design 85, 1086–1097.

Petrides, D., Carmichael, D., Siletti, C., Koulouris, A., 2014. Biopharmaceutical process optimization with simulation and scheduling tools. Bioengineering 1, 154–187.

Rabideau, G., Knight, R., Chien, S., Fukunaga, A., Govindjee, A., 1999. Iterative repair planning for spacecraft operations using the aspen system, in: Artificial Intelligence, Robotics and Automation in Space, p. 99.

Saint-Guillain, M., 2019. Robust operations management on mars, in: Proceedings of the International Conference on Automated Planning and Scheduling, pp. 368–376.

Saint-Guillain, M., Vaquero, T., Chien, S., Agrawal, J., Abrahams, J., 2021. Probabilistic temporal networks with ordinary distributions: Theory, robustness and expected utility. Journal of Artificial Intelligence Research 71, 1091–1136.

Smith, S.F., Lassila, O., 1993. Configurable systems for reactive production management., in: Knowledge-Based Reactive Scheduling, Citeseer. pp. 93–106.

Smith, S.F., Lassila, O., Becker, M., 1996. Configurable, mixed-initiative systems for planning and scheduling. Advanced Planning Technology , 235–241.

Stadtler, H., Stadtler, H., Kilger, C., Kilger, C., Meyr, H., Meyr, H., 2015. Supply chain management and advanced planning: concepts, models, software, and case studies. Springer.

Steel, R., Niézette, M., Cesta, A., Fratini, S., Oddi, A., Cortellessa, G., Rasconi, R., Verfaillie, G., Pralet, C., Lavagna, M., et al., 2009. Advanced planning and scheduling initiative mrspock aims for xmas. Language 1050, 3.

Turner, R.J., Huemann, M., Anbari, F.T., Bredillet, C.N., 2010. Perspectives on projects. Routledge.

Yelamanchili, A., Agrawal, J., Chien, S., Biehl, J., Connell, A., Guduri, U., Hazelrig, J., Maxwell, K., Steadman, K., Towey, S., 2020. Ground-based automated scheduling for the mars 2020 rover .