

Flexibility and its relationship to Decoupling in Simple Temporal Problems

Abstract

In this paper we concentrate on finding a suitable metric to determine the flexibility of a Simple Temporal Problem (STP). After reviewing some flexibility metrics that have been proposed, we conclude that these metrics fail to capture the correlation between events specified in the STP, resulting in an overestimation of the available flexibility in the system. We propose to use an intuitively more acceptable flexibility metric based upon uncorrelated time-intervals for the allowable starting times of events in an STP. This metric is shown to be computable in low-polynomial time. As a byproduct of the flexibility computation, we get a decomposition of the STN almost for free: for every possible k -partitioning of the event space, a decomposition can be computed in $O(k)$ -time. Even more importantly, we show that contrary to popular belief, such a decomposition does not affect the flexibility of the original STP.

1 Introduction

Scheduling problems occur in many diverse application domains such as transportation, process industry, health and education. Intuitively, a scheduling problem arises if there is a set of temporal events whose (starting) times are subject to a number of temporal constraints and we have to determine these starting times such that all constraints are satisfied. Usually, we are not satisfied with determining a single schedule, i.e., an assignment σ determining for every event t an exact starting time $\sigma(t)$. First of all, due to outside disturbances (e.g., unforeseen delays of preceding events), it may not be possible to start a certain event t at the predetermined time. Then it is essential for the scheduling agent to know how the current schedule can be maintained or needs to be adapted. Since during execution there is not much time for rescheduling, the possibilities for alternative starting times should be computable in a very short time or, even better, known in advance. Therefore, instead of offering a fixed schedule as a solution, one would like to compute an *interval schedule*, that is, an assignment of *time intervals* to variables, such that for every event t one can freely choose a specific time point in its interval. Such an interval for event t indicates the *flexibility*

we have w.r.t. starting t . The sum of such flexibilities then indicates the flexibility of the interval schedule and could be used to measure its quality in this respect.

Another reason to opt for scheduling with time intervals occurs when the scheduling problem is distributed among *multiple actors* who are involved in solving the overall scheduling problem. Typically, such agents want to schedule their part of the problem independently of the others, based on their own preferences and/or private objectives. One method that has been applied to fairly general scheduling formalisms is known as *temporal decoupling* [Hunsberger, 2002]. Here the original scheduling problem is decomposed into a number of subproblems, such that any solution (i.e., schedule) to a subproblem can be merged with any solution of the other subproblems, to constitute a feasible schedule of the total problem. As has been pointed out by various authors [Hunsberger, 2002; Planken *et al.*, 2010; Boerkoel and Durfee, 2011; Brambilla, 2010], when using existing flexibility metrics, applying temporal decoupling may affect the total flexibility of the system. The question then arises, how we could minimise this flexibility loss due to temporal decoupling. Hence, we would also like to (*i*) optimize the decoupling with respect to flexibility and, if necessary, (*ii*) investigate the loss of flexibility due to decoupling.

Our approach to addressing these problems is the following. We will concentrate on Simple Temporal Planning (STP) problems [Dechter, 2003] as the main representation mechanism for (temporal) scheduling problems. After a discussion of existing metrics to characterise the flexibility of STPs, we show that these metrics are not able to capture correlated time events, and generally overestimate the flexibility of scheduling systems. Then we introduce a new flexibility metric by using a transformation from a given STP S into a new STP S' such that correlations between time points are resolved in S' . By using an LP-approach to finding a simple uncorrelated interval schedule for events in S' , we can obtain the flexibility of the original system S . We show that in case of STPs this flexibility metric can be computed in polynomial time and offers the possibility to adapt existing schedules in $O(1)$ time. Then we deal with the problem of finding an optimal temporal decoupling for a given STP. Surprisingly, the method we use to specify the flexibility of an STP allows us to compute a decoupling for an STP that does not induce any loss of flexibility due to decoupling and can be obtained in $O(k)$ -time

(if the flexibility of the system has already been computed), where k is the number of agents involved in the decoupling. This result shows that temporal decoupling can be computed both efficiently and without loss of flexibility.

2 Preliminaries

The Simple Temporal Problem [Dechter *et al.*, 1991; Dechter, 2003], offers a convenient framework for analyzing temporal aspects of scheduling problems, distinguishing between a set of temporal variables and linear constraints between them.

An instance of the Simple Temporal Problem (STP) is a pair $S = (T, C)$, where $T = \{t_0, t_1, \dots, t_n\}$ is a set of time-point variables (events) and C is a finite set of binary constraints on T , each constraint having the form $t_j - t_i \leq c_{ji}$, for some real number c_{ji} . Given S , we want to find a schedule for the events in T , that is a function $\sigma : T \rightarrow \mathbb{R}^{\geq 0}$, assigning a non-negative real value (time-points) to each event in T , such that all constraints in C are satisfied. If such a schedule exists, we say that the STP is consistent. Without loss of generality, in the remaining of the paper, we simply assume that an STP is consistent.¹

In order to be able to express absolute time constraints, the time-point t_0 , also denoted by z , is used. It represents a fixed reference point on the timeline, usually having the value 0.

Often we represent an STP $S = (T, C)$ as a directed labeled graph where the vertices represent events in T and for every constraint $t_j - t_i \leq c_{ji}$ in C there is a directed edge (t_i, t_j) labeled by c_{ji} . The graph associated to an STP is called a Simple Temporal Network (STN). In this paper, we will use both representations interchangeably.

Example 1. In Figure 1 two STNs S_1 and S_2 are given, both representing an STP with three time events (including the time reference point z). From these graphs we can derive the specification of the associated STPs. For example, the STP associated to S_1 contains a constraint $t_2 - t_1 \leq 5$ and a constraint $t_1 - t_2 \leq -5$, while the STP belonging to S_2 contains a constraint $t_2 - t_1 \leq 0$ and a constraint $t_1 - t_2 \leq 5$.

The graph representation is very helpful in finding constraints implied by the set C of constraints, interpreting the weights c_{ji} on the arcs as lengths of the path from t_i to t_j . For example, if C contains the constraints $t_j - t_i \leq 10$ and $t_k - t_j \leq 15$, then there exists a path from t_i via t_j to t_k having a length $10 + 15 = 25$. This implies $t_k - t_i = (t_k - t_j) + (t_j - t_i) \leq 25$ is an implied constraint, even though C may not have contained a constraint limiting $t_k - t_i$. Using a shortest path interpretation of the STN $S = (T, C)$, there is an efficient method to find all *tightest* constraints implied by C . Basically, these tightest constraints can be found by searching for all shortest paths between the time points in T by e.g. applying Floyd and Warshalls all-pairs shortest paths algorithm [Floyd, 1962]. These tightest constraints are represented as the elements of the $n \times n$ distance matrix D_S , containing for every pair of time-point variables t_i and t_j the length of the *shortest path* in the distance graph between t_i and t_j . If D_S contains the entries $D_S[i, j]$ and

¹Consistency of an STP can be determined in low-order polynomial time.

$D_S[j, i]$, then $-D_S[j, i] \leq t_j - t_i \leq D_S[i, j]$ are the strongest constraints implied by C w.r.t. the temporal difference between t_j and t_i . The first row of D_S contains all entries $D_S[0, i]$ such that $t_i - z = t_i \leq D_S[0, i]$, so $D_S[0, i]$ is t_i 's latest starting time, $lst(t_i)$. Analogously, the first column contains entries $D_S[j, 0]$ such that $z - t_j \leq D_S[j, 0]$, i.e., $t_j - z = t_j \geq -D_S[j, 0]$, so $-D_S[j, 0]$ is t_j 's earliest starting time, $est(t_j)$. The distance matrix D_S for a given STN S can be computed in $O(n^3)$ time [Dechter, 2003]. Hence, using the STP machinery we can find earliest and latest starting times for events in $O(n^3)$ time.

The following proposition is a summary of two relevant properties that can be proven to hold for STNs and their schedules:

Proposition 1. Let $S = (T, C)$ be an STN and D_S its distance matrix. For $i = 1, \dots, n$, let $est(t_i) = D_S[0, i]$ and $lst(t_i) = -D_S[i, 0]$. Then, for every schedule σ for S , and every $t \in T$, it holds that $\sigma(t) \in [est(t), lst(t)]$. Moreover, given any $t \in T$ and $v \in [est(t), lst(t)]$, there exists² a schedule σ for S such that $\sigma(t) = v$.

Since we want to solve scheduling problems using STNs, we assume that for each event t there is a *finite horizon* for scheduling it. That means that for each STN $S = (T, C)$ we assume that there exists a finite constant h_S such that for all $t \in T$ it holds that $t \leq h_S$ can be derived from C . This avoids the use of unbounded time intervals (e.g. $t_i - t_j \leq \infty$).

3 Flexibility of an STN

Intuitively, the flexibility of an STN refers to the amount of freedom we have in assigning times to events in T . We defined a fixed schedule σ for an STN $S = (T, C)$ as an assignment $\sigma : T \rightarrow \mathbb{R}^{\geq 0}$ satisfying all the constraints in C . Such a schedule, however, is just an arbitrary fixed assignment among a large set of possible alternative assignments also satisfying the constraints. Therefore a single fixed schedule does not offer any indication of the flexibility of the STN it has been derived from.

Using earliest and latest starting times

If we want flexibility to indicate our freedom of choice, then one option for defining the flexibility for scheduling event t is to use the difference $lst(t) - est(t) = flex_N(t)$. The total flexibility $flex_N(S)$ of S then could be defined as

$$flex_N(S) = \sum_{t \in T} flex_N(t).$$

Obviously, $flex_N(S)$ can be computed in $O(n^3)$ time. Moreover, for each individual t , $flex_N(t)$ gives an exact measure of flexibility of event t , since, by Proposition 1, for every choice of a value $v \in [est(t), lst(t)]$ there exists a schedule σ for S such that $\sigma(t) = v$. However, although for every individual event, $flex_N(t)$ offers a precise measure of its flexibility, using the sum $flex_N(S)$ of these flexibilities has a serious disadvantage, due to *correlations* that might exist between starting times of events. To see this, consider the following example:

²In fact it can be shown that there exists a simple backtrack-free polynomial-time algorithm to construct an arbitrary schedule σ for S .

Example 2. Consider the STNs S_1 and S_2 (see Figure 1) where S_1 specifies the *concurrent* execution of $k = 3$ events within 5 time units, while S_2 specifies the *sequential* execution of $k = 3$ consecutive events, i.e., t_2 should not occur before t_1 , while t_3 should not occur before t_2 and all should occur within 5 time units.

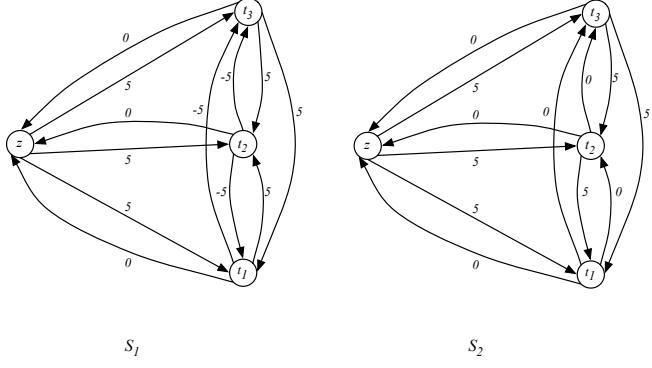


Figure 1: Two STNs each specifying three events t_1 , t_2 and t_3 in the interval $[0, 5]$. Here, S_1 specifies the *concurrent* execution of these events in the interval $[0, 5]$, while S_2 encodes their *sequential* execution $t_3 \leq t_2 \leq t_1$.

Note that both in S_1 and in S_2 , for each $i \in \{1, 2, 3\}$, $\text{est}(t_i) = 0$ and $\text{lst}(t_i) = 5$. Hence,

$$\text{flex}_N(S_1) = 3 \times 5 = 15 = \text{flex}_N(S_2).$$

So, according to the measure flex_N , S_1 and S_2 have the same flexibility, which, of course, is counterintuitive: For each event t_i in S_1 we can determine its starting time between 0 and 5 independently from the other events. So, the total flexibility, expressed as the amount of choice we have in selecting a starting time for an event equals $5 \times 3 = 15$. The choice for the starting time of an event t_{i+1} in S_2 , however, may influence the choice of the starting time of the next event t_i . More precisely, if (v_3, v_2, v_1) are the latest starting times for the 3 events in S_2 , it is clear that $0 \leq v_3 \leq v_2 \leq v_1 = 5$. Hence, the flexibility of t_3 is $v_3 - 0$, the flexibility of t_2 is $v_2 - v_3$ and the flexibility of t_1 is $5 - v_2$. So, the total flexibility of S_2 equals

$$\text{flex}(S_2) = v_3 + (v_2 - v_3) + (5 - v_2) = 5.$$

Therefore, the flexibility of S_2 should come out as only $\frac{1}{3}$ of the flexibility of S_1 . Generalising S_1 and S_2 to systems containing k concurrent and sequential events, respectively, we see that the freedom to schedule in S_1 equals $\text{flex}(S_1) = 5 \times k$, while the flexibility of S_2 is independent of k and remains $\text{flex}(S_2) = 5$. Hence, if k goes to infinity, the ratio of the real flexibilities of S_1 and S_2 goes to infinity:

$$\lim_{k \rightarrow \infty} \frac{\text{flex}(S_1)}{\text{flex}(S_2)} = \lim_{k \rightarrow \infty} \frac{5 \times k}{5} = \infty.$$

while their ratio according to flex_N is exactly 1:

$$\lim_{k \rightarrow \infty} \frac{\text{flex}_N(S_1)}{\text{flex}_N(S_2)} = \frac{5 \times k}{5 \times k} = 1.$$

We conclude that the flex_N metric seriously overestimates the flexibility of STNs similar to S_2 .

Using upper and lower bounds for every pair of events

One reason flex_N fails is that it is not able to deal with correlations between timed events. To overcome this disadvantage, we have to incorporate information about the *correlation* between two events as has been proposed in the definition of flexibility by Hunsberger [Hunsberger, 2002] and others [Pollicella et al., 2007]. Hunsberger defined his flexibility metric, which we will denote by $\text{flex}_H(S)$, by taking into account not only the apparent flexibility $\text{lst}(t) - \text{est}(t)$ of every event t , but also the flexibility between pairs of events t_i and t_j . Since the distance matrix D_S contains the entries $D_S[i, j]$ and $D_S[j, i]$ such that $-D_S[j, i] \leq t_j - t_i \leq D_S[i, j]$ are the strongest constraints implied by C w.r.t. the temporal difference $t_j - t_i$, we can define the flexibility associated with this difference as the length of the corresponding interval³:

$$\text{flex}_H(t_i, t_j) = D_S[i, j] + D_S[j, i].$$

Now the flexibility of the system S can be defined by taking $\text{flex}_N(S)$ and adding to it the sum of all these flexibilities for every distinct pair of time-points,⁴ i.e.,

$$\text{flex}_H(S) = \text{flex}_N(S) + \sum_{i=1}^n \sum_{j>i} \text{flex}_H(t_i, t_j).$$

If there is a correlation between two events t_i and t_j , this will result in a lower value of $D_S[i, j] + D_S[j, i]$. Hence, correlation among time events will have an influence on flex_H . Although this metric is an improvement w.r.t. the naive metric flex_N , the next example shows that flex_H also has its shortcomings.

Example 3. Consider again the STNs S_1 and S_2 in Figure 1. According to flex_H , the flexibility of S_1 equals $3 \times 5 + 3 \times 10 = 45$, while the flexibility of S_2 equals $3 \times 5 + 3 \times 5 = 30$. So, according to the Hunsberger metric, S_2 would have a flexibility equal to $\frac{2}{3}$ of the flexibility of S_1 . This seems to be an improvement with respect to the previous flexibility metric flex_N assigning the same flexibility to S_1 and S_2 . However, generalising S_1 and S_2 to systems containing k concurrent and sequential events, respectively, we see that

$$\text{flex}_H(S_1) = 5 \times k + \binom{k}{2} \times 10 = 5 \times k^2,$$

while

$$\text{flex}_H(S_2) = 5 \times k + \binom{k}{2} \times 5 = 5 \times \frac{k^2 - k}{2}.$$

This means that for larger values of k , the ratio between $\text{flex}_H(S_1)$ and $\text{flex}_H(S_2)$ converges to 2:

$$\lim_{k \rightarrow \infty} \frac{\text{flex}_H(S_1)}{\text{flex}_H(S_2)} = \lim_{k \rightarrow \infty} \frac{5 \times k^2}{5 \times \frac{k^2 - k}{2}} = 2,$$

while the ratio of the real flexibilities of S_1 and S_2 (see the previous example) goes to *infinity* if k goes to infinity. So, this metric is also not able to capture the correlations between the flexibilities of events in a satisfactory way.

³Although Hunsberger defined the inverse of flexibility, i.e., *rigidity* of an STN, that is not relevant for our discussion at hand, since we assume a finite time horizon for an STN.

⁴Actually, Hunsberger took the square root of the sum of the squares of these flexibility measures per pair, divided by $(n \times (n - 1)/2)$, the total number of different pairs of time-points.

3.1 Towards a suitable flexibility metric

Analysing the failure of both metrics $\text{flex}_N(S)$ and $\text{flex}_H(S)$ to capture the real flexibility of an STN S quickly reveals the main shortcoming of both of them: they do not fully capture the correlations between the allowed schedule-time intervals for the events t , if such correlations do exist. Considering again the system S_1 (see Figure 1), it can be observed that the real flexibility can be computed by taking the sum $\text{flex}(t) = \text{flex}_N(t)$ of each event. This latter flexibility is simply determined as the difference between the latest starting time and the earliest starting time of such an event. But, as we have seen before, as soon as the events are correlated (as in the STN S_2), this no longer holds, because the starting time that can be chosen for an event t is dependent upon the starting times that were chosen for other events $t' \in T$.

So, in order to compute the flexibility of a system S , the flexibility measure should be based upon *uncorrelated* starting time intervals.

Definition 1. Given an STN $S = (T, C)$, a set $I_S = \{[l_t, u_t] \mid t \in T\}$ of (non-empty) intervals for the variables in T is uncorrelated iff for every $t \in T$ and every $v_t \in [l_t, u_t]$, the assignment σ , defined by $\sigma(t) = v_t$, is a schedule for S .

Then, to compute the real flexibility of a system $S = (T, C)$, we have to determine

1. a set I_S of uncorrelated starting time intervals $[l_t, u_t]$, such that
2. the sum of the interval lengths $\sum_{t \in T} (u_t - l_t)$ is maximal.

The idea we will now pursue is to transform an STN S to an STN S' in which new variables t_i^- and t_i^+ occur for every $t_i \in T$. The intuitive meaning of these variables is that they constitute the earliest and latest starting times of t_i in a set I_S of uncorrelated intervals.

Next, we determine the maximum flexibility of S using an LP-formulation to find a maximal solution satisfying both requirements.

Let's start with the first step. In order to ensure that the variables t_i^- and t_i^+ correspond to their intended meaning, the following proposition shows how to obtain a set of uncorrelated intervals. Intuitively, this proposition states that if there exists a constraint on some $t_j - t_i$ in C , then the intervals for these events should be constrained so that, within their respective intervals, the events can not be scheduled so far apart as to violate the constraint:

Proposition 2. Let $S = (T, C)$ be an STN. A set $I_S = \{[l_{t_i}, u_{t_i}] \mid t_i \in T\}$ of intervals for the variables in T is uncorrelated if for every pair $(t_i, t_j) \in T^2$, it holds that $t_j - t_i \leq c_{ji} \in C$ implies $u_{t_j} - l_{t_i} \leq c_{ji}$.

Proof. Suppose the implication holds for every pair of variables. We have to show that I_S is uncorrelated. So, for every $t \in T$, choose an arbitrary value $v_t \in [l_t, u_t]$ and let σ be defined by $\sigma(t) = v_t$, for all $t \in T$. We now prove that σ is a schedule for S . Suppose σ is not a schedule, meaning that σ violates some constraint in C . Then there exists a constraint $t_j - t_i \leq c_{ji} \in C$ while $\sigma(t_j) - \sigma(t_i) > c_{ji}$. But this implies that $c_{ji} < v_{t_j} - v_{t_i} < u_{t_j} - l_{t_i}$, hence $u_{t_j} - l_{t_i} > c_{ji}$, contradicting the fact that $u_{t_j} - l_{t_i} \leq c_{ji}$ (as we have assumed to be implied by $t_j - t_i \leq c_{ji} \in C$). \square

We will now show how to construct an STN in which the interval-endpoints are variables and the required constraints are imposed, so that, if we solve this derived STN, we are guaranteed suitable interval-endpoints:

Proposition 3. Let $S = (T, C)$ be an STN. Consider the STN $S' = (T', C')$ derived from S as follows:

1. $T' = \{t^+, t^- \mid t \in T - \{z\}\} \cup \{z\}$;
2. $C' = \{t^+ - t'^- \leq c \mid \forall (t - t' \leq c) \in C\} \cup \{z - t^- \leq c \mid z - t \leq c \in C\} \cup \{t^+ - z \leq c \mid t - z \leq c \in C\}$.

Then, for every solution σ of S' , the set $\{[\sigma(t^-), \sigma(t^+)] \mid t \in T\}$ is an uncorrelated set of solution intervals for S .

Proof. If C contains a constraint of the form $t_j - t_i \leq c_{ji}$, then by the specification of the first group of constraints in C' , any assignment $\sigma : T' \rightarrow \mathbb{R}^{>0}$ that satisfies these constraints, also satisfies $\sigma(t_j^+) - \sigma(t_i^-) \leq c_{ji}$, as required in Proposition 2. \square

Example 4. Consider the STN S_2 presented in Example 1. Let STN $S'_2 = (T', C')$ be the transformed STN where

1. $T' = \{z, t_1^-, t_2^-, t_3^-, t_1^+, t_2^+, t_3^+\}$ and
2. $C' = \{t_i^+ - t_j^- \leq 5 \mid 1 \leq i < j \leq 3\} \cup \{t_i^+ - t_j^- \leq 0 \mid 1 \leq j < i \leq 3\} \cup \{z - t_i^- \leq 0 \mid i = 1, 2, 3\} \cup \{t_i^+ - z \leq 5 \mid i = 1, 2, 3\}$

One possible solution is $\sigma(t_1^+) = 4$, $\sigma(t_2^+) = 3$ and $\sigma(t_3^+) = 1$, while $\sigma(t_1^-) = 3$, $\sigma(t_2^-) = 1$ and $\sigma(t_3^-) = 0$. This results in the following set of uncorrelated solution intervals: $\sigma_I(t_3) = [0, 1]$, $\sigma_I(t_2) = [1, 3]$, $\sigma_I(t_1) = [3, 4]$.

Although our transformation of S into S' now provides an uncorrelated set of intervals, it is still insufficient to determine the flexibility of S , since the sum of the sizes of these intervals does not need to be maximal. To obtain the flexibility measure, we have to select those solutions σ for S' that maximise the sum of the sizes of the intervals. It is easy to see that the following LP formulation will solve this problem:

Theorem 1. Let $S = (T, C)$ be an STN. Then $\text{flex}(S)$ can be computed by solving the following Linear Program:

$$\begin{aligned} \max \quad & \sum_{t \in T} (t^+ - t^-) \\ \text{subject to:} \quad & t^- \leq t^+ \quad \forall t \\ & t^+ - t'^- \leq c \quad \forall (t - t' \leq c) \in C \end{aligned} \quad (1)$$

As an easy consequence of these results we have the following corollary.

Corollary 1. Given an STN $S = (T, C)$, the flexibility metric flex determines an assignment $\text{flex}(t) = [t^-, t^+]$, such that for every $t \in T$ and every $v_t \in [t^-, t^+]$, the assignment $\sigma(t) = v_t$ is a fixed-time schedule for S , satisfying all constraints in C .

Example 5. Applying the flexibility metric flex to the STNs S_1 and S_2 results in the following flex functions:

1. For S_1 , we have $\text{flex}(t_i) = [0, 5]$ for $i = 1, 2, 3$ and $\text{flex}(S_1) = 3 \times 5 = 15$. \square

2. For S_2 , we have $\text{flex}(S_2) = 5$; one possible flex function is $\text{flex}(t_1) = [0, 5]$, $\text{flex}(t_2) = \text{flex}(t_3) = [5, 5]$.

So as intended, this flexibility function will assign 3×5 flexibility units to S_1 , while assigning 5 units to S_2 .

4 Decoupling without loss of flexibility

STNs might also be used in a distributive scheduling context (see [Hunsberger, 2002; Boerkoel and Durfee, 2011; 2012; Brambilla, 2010]). Suppose that the events in T are distributed over k agents A_1, \dots, A_k such that each agent is responsible for scheduling a disjoint subset $T_k \subseteq T - \{z\}$ of events. These agents act as distributed autonomous schedulers, so each agent A_i wants to determine its own schedule $\sigma_i : T_i \rightarrow \mathbb{R}^{\geq 0}$ for the STP $S_i = (T_i \cup \{z\}, C'_i)$, where T_i is the subset of tasks (time-points) assigned to A_i and $C'_i = T_i^2 \cap C$ is a set of constraints restricted to the events in T_i . In allowing agents to schedule autonomously, we have to ensure that no matter what choices the individual agents make in composing their individual schedule, *all* constraints in C are satisfied. That is, the merge $\sigma = \bigcup_{i=1}^k \sigma_i$ of these individual schedules should always constitute a valid total schedule to the original problem. This problem is known as the so-called *Temporal Decoupling Problem* [Hunsberger, 2002].

Definition 2. Let $S = (T, C)$ be a consistent STN. Suppose that $T \setminus \{z\} = \{T_i\}_{i=1}^k$ is partitioned in k subsets T_i . Then the temporal decoupling problem is to find k STPs $S_i = (T_i \cup \{z\}, C'_i)$ such that, whenever $\sigma_1, \dots, \sigma_k$ are independently constructed schedules for the individual STPs S_1, \dots, S_k , respectively, their merge $\sigma = \bigcup_{i=1}^k \sigma_i$ is a schedule for the original STP S .

To achieve an arbitrary decoupling of an STN $S = (T, C)$ where T is partitioned into blocks T_i , we have to pay attention to constraints involving events t_i and t_j that belong to different agents A_m and A_n . The problem is that agents A_m and A_n may choose schedules σ_m and σ_n , respectively, that violate inter-agent constraint $t_i - t_j \leq c_{ij}$, by setting $\sigma_m(t_i) - \sigma_n(t_j) > c_{ij}$. Hunsberger [Hunsberger, 2002] gave an elegant solution to this temporal decoupling problem by proposing an iterative procedure that basically makes each *inter-agent* constraint $t_i - t_j \leq D_S[j, i]$ obsolete by tightening *intra-agent* constraints $t_i - z \leq D_S[0, i]$ and $z - t_j \leq D_S[j, 0]$. These tightenings consist in selecting values $\delta_{i,1}$ and $\delta_{i,2}$ such that

1. $t_i - z \leq \delta_{i,1} \leq D_S[0, i]$,
2. $z - t_j \leq \delta_{i,2} \leq D_S[j, 0]$, and
3. $\delta_{i,1} + \delta_{i,2} \leq D_S[i, j]$.

If we add the new constraints $t_i - z \leq \delta_{i,1} \leq D_S[0, i]$ and $z - t_j \leq \delta_{i,2} \leq D_S[j, 0]$ to S while making sure that the resulting system remains consistent, the effect of adding these constraints is that an inter-agent constraint $t_i - t_j = (t_i - z) + (z - t_j) \leq \delta_{i,1} + \delta_{i,2} \leq D_S[i, j]$ is now *implied* by the new, tighter, intra-agent constraints. Hence, the inter-agent constraint can be removed from the updated STP. Applying this procedure for every non-implied inter-agent constraint ensures that all inter-agent constraints are implied by

intra-agent constraints and the resulting system is a temporal decoupling.

Example 6. Suppose that in the STNs S_1 and S_2 from our earlier examples, we have agents A_1, A_2 and A_3 , where A_i is responsible for scheduling t_i . In S_1 , every inter-agent constraint $-5 \leq t_i - t_j \leq 5$ is already implied by the intra-agent constraints $0 \leq t_i - z \leq 5$ and $0 \leq t_j - z \leq 5$, so decoupling is already achieved. In S_2 , however, we can decouple by tightening the intra-agent constraints to (for example) $0 \leq t_3 - z \leq 2$, $3 \leq t_2 - z \leq 4$ and $4 \leq t_1 - z \leq 5$, so that every inter-agent constraint is implied.

An arbitrary decoupling, however, is not always what we want, since the added constraints may limit the resulting flexibilities of the subsystems S_i . Hence, the sum $\sum_{i=1}^k \text{flex}(S_i)$ of the flexibilities of the subsystems S_i could be considerably less than the flexibility $\text{flex}(S)$ of the original system. It is a popular belief in the research community that an optimal decoupling i.e., a decoupling that would maximize the flexibilities of the subsystems, would still suffer from some flexibility loss. In the literature we are aware of (see e.g. [Hunsberger, 2002; Boerkoel and Durfee, 2011; 2012]) several experiments have been performed to measure this loss of flexibility. However, the metrics used all have been based on the flex_H metric discussed above. Instead of using this metric we try to find a decoupling that maximizes the sum of the flexibilities of the induced subsystems S_i , and we consider it an *optimal decoupling* if the ratio $(\sum_{i=1}^k \text{flex}(S_i)) / \text{flex}(S)$ is maximal.

Contrary to the popular beliefs in the research community we will show now that

1. an optimal decoupling can be obtained in $O(k)$ -time using the flexibility metric we discussed before,
2. this optimal decoupling need not affect the flexibility of the system, i.e., we have $\sum_{i=1}^k \text{flex}(S_i) = \text{flex}(S)$.

Proposition 4. Let $\{S_i\}_{i=1}^k$ be an optimal decoupling of an STN S . Then $\sum_{i=1}^k \text{flex}(S_i) = \text{flex}(S)$.

Proof. Consider the set of intervals $\{[l_t, u_t] : t \in T\}$ occurring as solutions of the LP (1). Given the set $A = \{A_i\}_{i=1}^k$ of k agents, let $C_{\text{inter}} \subset C$ be the set of all inter-agent constraints. For every inter-agent constraint $t - t' \leq c$, where t occurs in S_i and t' occurs in S_j , add the constraint $t - z \leq u_t$ to C_i and add $z - t' \leq -l_{t'}$ to C_j . We show the following.

Claim 1: The resulting systems $\{S_i\}_{i=1}^k$ constitute a decoupling of S ;

Claim 2: The sum of the flexibilities $\text{flex}(S_i)$ of the systems S_i in the decoupling $\{S_i\}_{i=1}^k$ equals $\text{flex}(S)$.

Ad Claim 1. Suppose, on the contrary, that the merge σ of some combination of individual schedules σ_i , violates a constraint in C . This must be an inter-agent constraint $t - t' \leq c$ where t occurs in S_i and t' occurs in S_j ($i \neq j$). So $\sigma_i(t) - \sigma_j(t') > c$, but since σ_i satisfies all constraints in C_i , it follows that $\sigma_i(t) \leq u_t$, and similarly, $\sigma_j(t') \geq l_{t'}$. Therefore, $\sigma_i(t) - \sigma_j(t') \leq u_t - l_{t'}$. Since the constraint $t^+ - t'^- \leq c$ occurs in the LP, we also know that $u_t - l_{t'} \leq c$, so $\sigma_i(t) - \sigma_j(t') \leq c$, which contradicts $\sigma_i(t) - \sigma_j(t') > c$.

Ad Claim 2. By definition, we have $\sum_{i=1}^k \text{flex}(S_i) \leq \text{flex}(S)$. Let I_S be a set of (uncorrelated) intervals $[t_t, u_t]$ realizing maximum flexibility. For every subsystem $S_i = (T_i, C_i)$, consider the set I_{S_i} , being the restriction of I_S to T_i . It is easy to see that each such set of intervals I_{S_i} constitutes a set of uncorrelated intervals, realizing maximum flexibility for S_i . But then $\sum_i \text{flex}(S_i) = \text{flex}(S)$. Hence, there exists at least one decoupling realizing a total flexibility equal to the original flexibility $\text{flex}(S)$. \square

Example 7. Consider Example 6. A decoupling for S_2 realizing maximum flexibility can be created by adding the following decoupling constraints:

$$t_2 - z \leq 5; \quad z - t_2 \leq -5; \quad t_3 - z \leq 5; \quad z - t_3 \leq -5.$$

The sum of flexibilities of the subsystems equals the flexibility of the original system.

Example 8. To illustrate the difference between the current approach and previous approaches to flexibility and decoupling, we apply our method to the following STN (see Figure 2 adapted from [Boerkoel, 2012]): The STN models the activities and their constraints for three agents Chris, Ann and Bill. The purpose of this example is to show how to apply decoupling in a real-life situation, where each of the agents should have enough freedom to schedule its own activities. The flexibility of the STN S according to the flex_N metric is 780 (minutes). Due to several dependencies this is a serious overestimation of the real flexibility of the system, which according to our metric equals $\text{flex}(S) = 180$ (minutes). In Figure 3, the result of decoupling the STN according to Boerkoel has been depicted. If this decoupling would have been optimal, we would expect its total flexibility to be equal to 180. However, the sum of the flexibilities of the three subsystems equals 135. Hence, we conclude that the decoupling applied can be improved by using our flexibility approach.

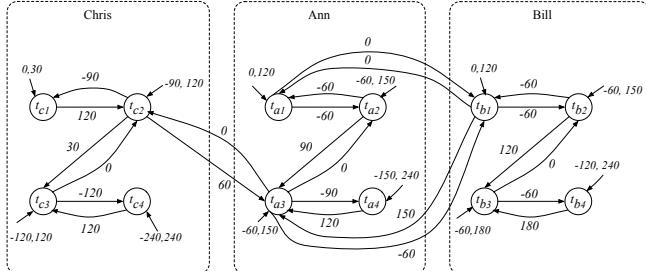


Figure 2: An STN adapted from [Boerkoel, 2012]. The arrows without source labeled with two numbers $[x, y]$ pointing towards an event t represent the weights of the constraints $z - t \leq x$ and $t - z \leq y$.

5 Conclusion and discussion

In this paper we have discussed an efficient method to determine the flexibility of a schedule for an STN. The advantage of the flexibility metric we propose is twofold: First of all, this metric takes into account the correlations between time events unlike previous methods that have been proposed in

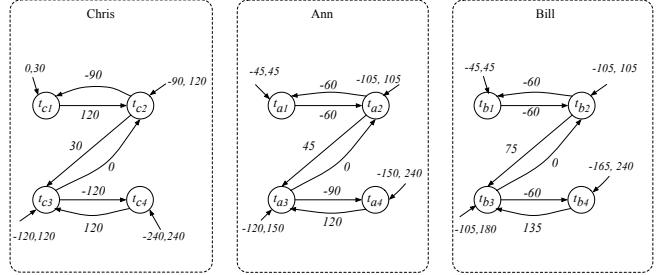


Figure 3: The decoupled STN from the previous figure (adapted from [Boerkoel, 2012]).

the literature. As a consequence, our metric does not suffer from overestimations of the flexibility in an STN. Secondly, using this improved flexibility metric allows us to disprove a rather widespread belief in the research community, i.e., that temporal decoupling of an STN necessarily results in reduced flexibility: We showed that once the flexibility of an STN is determined, it can be used to decide upon a temporal decoupling that does not affect the (total) flexibility of the system.

The total flexibility of a system, however, can be obtained in many, sometimes even quite unfair, ways. For example, preserving the total flexibility of the system could mean that a decoupling heavily favors some specific agents by giving them a lot of flexibility, while harming other agents by taking away their flexibility. To prevent this, we need to come up with fair flexibility distribution schemes as an objective to grant every agent an almost equal portion of flexibility. We could propose more *fair* distributions of flexibility and evaluate their performance by comparing it to the maximally obtainable flexibility as we have proposed in our paper to establish the *price of fairness*.

Another question we did not deal with in this paper is the question how we should distribute flexibility over the events in order to ensure that a given interval schedule as we have proposed in this paper could enhance the robustness of a schedule. For example, if certain events are considered as critical or more susceptible to delays, we could try to provide them with a larger scheduling interval. Research in this direction also should pay attention to the dependencies of events upon other events (as expressed by the STN) to minimize scheduling violations. Together with extensions of the STN framework with *stochastic* interpretations of the upper bounds c_{ij} on the constraints, this would enable stochastic interpretations of the total flexibility of an STN and encourage empirical research into optimal distributions of flexibility over the events in order to ensure maximal (expected) robustness of the resulting schedule.

Finally, note that temporal decoupling is now done off-line before the execution of the schedule. It would be quite interesting to come up with dynamic variants of decoupling where, based on the completion times of tasks already completed, the current decoupling is adapted to profit from flexibility in an even better way.

References

- [Boerkoel and Durfee, 2011] J.C. Boerkoel and E.H. Durfee. Distributed algorithms for solving the multiagent temporal decoupling problem. In *Proceedings AAMAS*, 2011.
- [Boerkoel and Durfee, 2012] J.C. Boerkoel and E.H. Durfee. A distributed approach to summarizing spaces of multiagent schedules. In *Proceedings AAAI*, 2012.
- [Boerkoel, 2012] J.C. Boerkoel. *Distributed Approaches for Solving Constraint-based Multiagent Scheduling Problems*. PhD thesis, University of Michigan, 2012.
- [Brambilla, 2010] A. Brambilla. *Artificial Intelligence in Space Systems: Coordination Through Problem Decoupling in Multi Agent Planning for Space Systems*. Lambert Academic Publishing, 2010.
- [Dechter *et al.*, 1991] R. Dechter, I. Meiri, and J. Pearl. Temporal constraint networks. *Artificial Intelligence*, 49:61–95, 1991.
- [Dechter, 2003] R. Dechter. *Constraint processing*. Morgan Kaufmann, 2003.
- [Floyd, 1962] R.W. Floyd. Algorithm 97: Shortest path. *Communications of the ACM*, 5(6):345, 1962.
- [Hunsberger, 2002] L. Hunsberger. Algorithms for a temporal decoupling problem in multi-agent planning. In *Proceedings AAAI*, 2002.
- [Planken *et al.*, 2010] L. Planken, M. de Weerdt, and C. Witteveen. Optimal temporal decoupling in multiagent systems. In *Proceedings AAMAS*, 2010.
- [Policella *et al.*, 2007] N. Policella, A. Cesta, A. Oddi, and S.F. Smith. From precedence constraint posting to partial order schedules: A CSP approach to robust scheduling. *AI Communications*, 20:163–180, 2007.