Using Constraint Networks on Timelines to Model and Solve Planning and Scheduling Problems

Cédric Pralet and Gérard Verfaillie
ONERA, Centre de Toulouse, DCSD
2 av. Edouard Belin
31400 Toulouse, France
cedric.pralet,gerard.verfaillie@onera.fr

Abstract

In the last decades, there has been an increasing interest in the connection between planning and constraint programming. Several approaches were used, leading to different forms of combination between the two domains. In this paper, we present a new framework, called Constraint Network on Timelines (CNT), to model and solve planning and scheduling problems. Basically, CNTs are a kind of dynamic CSPs, enhanced with special variables called dimension variables representing the initially unknown number of steps in a valid or optimal plan. We also present an algorithm and experimental results showing that the expressiveness of CNTs allows efficient models to be developed, and can lead to significant gains on problems taken from planning competitions.

Introduction

In the last decades, there has been an increasing interest in the connection between planning and Constraint Programming (CP). As already recognized in (Nareyek et al. 2005), this interest has led to three main kinds of combination between planning and CP.

1. CP can be used as a plug-in to solve efficiently subproblems generated during planning. This plug-in approach allows existing planners to be enhanced with CP techniques but does not exploit all the capabilities of CP.

2. In approaches inspired from (Kautz & Selman 1992), such as CPlan (van Beek & Chen 1999) or GP-CSP (Do & Kambhampati 2001), a CSP (Constraint Satisfaction Problem (Dechter 2003)) or a dynamic CSP (Mittal & Falkenhainer 1990) is built to solve the planning problem over a fixed horizon k, which is incremented if no plan is found. The size-bounded CSPs constructed are obtained from planning graphs (Blum & Furst 1997) or directly from STRIPS or PDDL representations (Fikes & Nilsson 1971) [McDermott 1998]. They contain variables representing the state and the actions at each step i ∈ [1..k], and constraints specifying the initial and goal states, action preconditions, and action effects. Other constraints may be added manually. This approach can be very efficient but as all variables are duplicated at each step, the CSPs built can become too large.

3. Other approaches tackle planning problems as a kind of dynamic CSP without fixing the horizon. This includes planners like CPT (Vidal & Geffner 2006), which represents the planning problem by a set of temporal variables associated with actions and action preconditions, and by a set of temporal constraints. The associated CSP is directly obtained from PDDL descriptions. It is dynamic in the sense that at each step of the search, variables and constraints are active or not. (Nareyek 2001) proposes another approach, which involves a CSP with a dynamic and constrained graphical structure. Another example is Constraint-based Attribute and Interval Planning (Frank & Jonson 2003), whose principle is to add to a current incomplete plan so-called intervals. The latter represents that some predicate holds over a time slot [t_s, t_e], and must satisfy some compatibility constraints.

In this paper, we propose a new generic constraint-based approach to model and solve planning and scheduling problems. This approach, called Constraint Network on Timelines (CNT), is included in the third category but covers all approaches in the second one. A CNT is a kind of dynamic CSP, in which the dynamic aspect comes from the explicit presence of dimension variables representing the possibly unknown number of steps in the plans sought. The presentation here differs from the first version of CNTs introduced in (Verfaillie, Pralet, & Lemaître 2008). The paper is organized as follows. We first present the CNT framework and compare it with existing approaches. The different modeling capabilities of CNTs are illustrated on some planning problem examples taken from international planning competitions (IPCs). We then present an algorithm to seek for plans and optimal plans from CNT representations, highlighting the special role of dimension variables. Last, we give experimental results on IPC problems, showing that using the modeling capabilities and the expressiveness of CNTs can lead to significant gains in computation time. In particular, some problems unsolved by existing optimal planners are solved optimally in a few seconds.

Constraint Networks on Timelines

In the following, [a..b] denotes the set of integers between a and b, and given a variable x, d(x) denotes its domain of values. In order to illustrate the CNT framework, we consider a space application example (Pralet & Verfaillie 2008).

Copyright © 2008, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.
A satellite must download a set of $N_O$ observations down to Earth over a time period [STA, END]. Over this time period, it goes through a set $\{SE_k, EE_k, k \in \{1..N_E\}\}$ of $N_E$ eclipse periods, and a set $\{SD_k, ED_k, k \in \{1..N_D\}\}$ of $N_D$ station visibility windows, during each of which a block of observations can be downloaded. Downloading observation $o \in \{1..N_O\}$ takes a duration $D_o$. At any time, the level of energy available on board is somewhere between $EN_{min}$ and $EN_{max}$. Its evolution depends on the power $P_d$ consumed by downloads, the power $P_{sol}$ consumed by the platform, and the power $P_{sol}$ produced by solar panels when the satellite is not in eclipse. The goal is to download all observations while respecting energy limitations. The initial level of energy, at time STA, is denoted $EN_{init}$.

To model this problem, we can first define a set of “classical” variables: for each observation $o$, we introduce one variable $na_o$ of domain $d(na_o) = [0..N_D]$ to represent the index $k$ of the download slot during which observation $o$ is downloaded (value 0 if $o$ is not downloaded), and for each download slot index $k \in \{1..N_D\}$, we introduce variables $sd_k$ and $ed_k$, of domain $d(sd_k) = d(ed_k) = [SD_k, ED_k]$ to represent respectively the start and end times of the download occurring during download slot $k$. To model the evolution of the energy level without fixing the number of steps in this evolution, we need to introduce a set of variables $en_i$, whose cardinality is unknown. This will be possible in the CNT framework thanks to the notion of timeline.

Definition 1 (Timeline) A timeline $tl$ is a pair $tl = (d(tl), h(tl))$ where $d(tl)$ is a set of values and $h(tl)$ is a variable whose domain of values $d(h(tl))$ is included in $\mathbb{N}$. $d(tl)$ is called the domain of $tl$ and $h(tl)$ its dimension (dimension is denoted $h$ like horizon).

Definition 2 (Variables associated with a timeline) Given an assignment $A$ of $h(tl)$, a timeline $tl = (d(tl), h(tl))$ defines a finite set of variables $V(tl, A) = \{tl_i \mid i \in \{1..A\}\}$, whose domain of values is $d(tl_i) = d(tl)$. This set is empty if $h(tl)$ takes value 0. In order to distinguish variables defined by a timeline from classical variables, variables in $V(tl, A)$ are called timeline-variables, or more shortly t-variables.

The maximal set of t-variables which may be defined by a timeline is $\{tl_i \mid i \in \{1..max(d(h(tl)))\}\}$. This set can be infinite if $d(h(tl))$ is not bounded. Among t-variables in this set, the ones in $\{tl_i \mid i \in \{1..min(d(h(tl)))\}\}$ are mandatory, because they exist whatever the timeline dimension is.

Definition 3 (Assignment of a set of timelines) Let $T$ be a set of timelines. An assignment $A_T$ of $T$ is the union of an assignment $A_{H}$ of all the dimension variables of the timelines in $T$, and of an assignment $A_{V}$ of all t-variables in $\bigcup_{tl \in T} V(tl, A_{H}[h(tl)])$, where $A_{H}[h(tl)]$ denotes the assignment of $h(tl)$ in $A_{H}$.

Let us illustrate these notions on the satellite example. We can first introduce one dimension variable $h$ of domain $d(h) = [1..\infty]$ to represent the number of important steps in the evolution of the level of energy. We use four timelines, $t = \{(STA, END), h\}$, $en = \{[EN_{min}, EN_{max}], h\}$, $ec = \{(0, 1), h\}$, and $dl = \{(0, 1), h\}$, to represent respectively the time associated with each step, the current level of energy, the current eclipse status, and the current download status. Given an assignment $A$ of $h$, these timelines induce a set of t-variables $\{tl_i \mid t \in \{t, en, ec, dl\}, i \in \{1..A\}\}$, where $tl_i$ represents the value of $tl$ at step $i$. A tabular representation of an assignment of the different timelines is given below. The first column means that at $t_1 = 0$, the energy level is $en_1 = 300$, the satellite is in eclipse ($ec_1 = 1$), and no download is performed ($dl_1 = 0$). At step 2, at $t_2 = 30$, the energy level has decreased to $en_2 = 270$, and a download is triggered ($dl_2 = 1$). The download ends at $t_3 = 48$ ($dl_3 = 0$). At $t_4 = 150$, the satellite is not in eclipse anymore ($ec_4 = 0$). And so on until step $h = 8$.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$</td>
<td>0</td>
<td>30</td>
<td>48</td>
<td>150</td>
<td>235</td>
<td>248</td>
<td>272</td>
<td>350</td>
</tr>
<tr>
<td>$en$</td>
<td>300</td>
<td>270</td>
<td>216</td>
<td>114</td>
<td>284</td>
<td>271</td>
<td>223</td>
<td>145</td>
</tr>
<tr>
<td>$ec$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$dl$</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Variables defined by timelines must usually satisfy some constraints. We therefore introduce the notion of constraint on timelines.

Definition 4 (Constraint) A classical CSP constraint $c$ is defined by a pair $(S(c), R(c))$ where $S(c)$ (the scope of $c$) is the finite set of variables over which the constraint holds, and $R(c)$ (the relation associated with $c$) is any explicit or implicit representation of the set of allowed combinations of values of the variables in $S(c)$.

Definition 5 (Constraint on timelines) A constraint on timelines is a triple $c = (S_V(c), S_T(c), fct(c))$ where $S_V(c)$ is a finite set of variables, $S_T(c)$ is a finite set of timelines, and $fct(c)$ is a function which associates a finite set of CSP constraints with each assignment of $A$ of the dimension variables of the timelines in $S_T(c)$. It is moreover assumed that the scope of each of the CSP constraints in $fct(c)(A)$ is included in $S_V(c) \cup \{tl_i \mid tl_i \in S_T(c), i \in \{1..A[h(tl)]\}\}$.

Given a timeline $tl$ and a variable $x$, an example of constraint on timelines is $c : \forall i \in \{1..h(tl)\}, tl_{i+1} \neq x + tl_i$. Implicitly, it is the triple $(S_V(c), S_T(c), fct(c))$ where $S_V(c) = \{x\}$, $S_T(c) = \{\}\$, and $fct(c)$ is the function which associates, with each assignment $A$ of $h(tl)$, the set of CSP constraints $\{tl_{i+1} \neq x+tl_i \mid i \in \{1..A-1\}\}$. Another example of constraint on timelines is: $\exists i \in \{2..h(tl)\}, tl_i = 0$. It corresponds to the triple $(\emptyset, \{tl\}, fct)$, where $fct$ associates, with each assignment $A$ of $h(tl)$, the CSP constraint $\exists i \in \{2..A\}, tl_i = 0$. In order to impose conditions on the final state of a timeline, constraints of the form $tl_{h(tl)} = g$ can be considered. They associate, with each assignment $A$ of $h(tl)$, the CSP constraint $tl_A = g$. We can also define constraints like $\text{alldifferent}(tl_i \mid i \in \{1..h\})$, which implicitly associates with each assignment $A$ of $h$ the classical CSP constraint $\text{alldifferent}(tl_i \mid i \in \{1..A\})$. Such constraints, usually called global constraints because they hold on a large number of variables, are interesting because they can be handled by dedicated powerful inference algorithms available in constraint programming tools.

To be more concrete, the satellite problem can be modeled by constraints listed below. Constraint $\text{alldifferent}$ defines the initial state. Constraint $\text{alldifferent}$ defines the end time of a download.
as its start time plus its duration. Constraints \( C \) and \( v \) define respectively when the satellite is in an eclipse/download state, \( C \) defines the evolution of the level of energy, and \( v \) asserts that the time \( t_{i+1} \) of step \( i + 1 \) is the minimum of all important time points strictly greater than \( t_i \). Last, \( o \) defines a condition on the final state and \( d \) defines the problem goal. Note that the domain of values associated with timelines also enforce constraints: e.g., the domain of values \( d(\epsilon) = [ENmin, ENmax] \) of timeline \( en \) imposes a constraint on the minimum level of energy.

\[
(\epsilon_1 = STA) \land (\epsilon_1 = ENinit) \quad (c_1)
\]

\[
\forall k \in \{1..N_D\}, ed_k = sd_k + \sum_{o \in \{1..N_O\}} |n_o = k| D_o \quad (c_2)
\]

\[
(\forall i \in \{1..h\}, \quad (\exists k \in \{1..N_E\}, SE_k \leq t_i < EE_k) \quad (c_3)
\]

\[
\forall i \in \{1..h\}, \quad (\exists k \in \{1..N_D\}, sd_k \leq t_i < ed_k) \quad (c_4)
\]

\[
t_{i+1} = \min\{t \in T_D \cup T_E \cup \{END\} \mid t > t_i\} \quad (c_5)
\]

\[
T_D = \cup_{k \in \{1..N_D\}} \{sd_k, ed_k\} \quad (c_6)
\]

\[
t_E = \cup_{k \in \{1..N_E\}} \{SE_k, EE_k\} \quad (c_7)
\]

\[
\forall o \in \{1..N_O\}, nd_o \neq 0 \quad (c_8)
\]

All notions defined previously are assembled in the notion of constraint network on timelines (CNT).

**Definition 6** (Constraint network on timelines) A constraint network on timelines \( \text{cnt} = (V, C_V, T, C_T) \), where \( V \) is a finite set of variables, \( C_V \) is a finite set of constraints whose scopes are included in \( V \), \( T \) is a finite set of timelines whose dimensions are included in \( V \), and \( C_T \) is a finite set of constraints on timelines \( (S_V, S_T, \text{fct}) \) such that \( S_V \subseteq V \) and \( S_T \subseteq T \).

The satellite download problem can be modeled by the CNT \( (V, C_V, T, C_T) \), where \( V = \{nd_o \mid o \in \{1..N_O\}\} \cup \{\cup_{k \in \{1..N_D\}} \{sd_k, ed_k\}\} \cup \{h\}, C_V = \{c_1, c_2, c_3\}, T = \{t, en, ec, dl\}, \text{and } C_T = \{c_1, c_3, c_4, c_5, c_6, c_7\}. \) The following figure gives a partial illustration of this CNT.

Among the various problems which can be formulated on CNTs, a useful one is to seek for a consistent assignment:

**Definition 7** (Consistent assignment of a CNT) A consistent assignment (a solution) of a constraint network on timelines \( \text{cnt} = (V, C_V, T, C_T) \) is an assignment of the variables in \( V \) and of the timelines in \( T \) such that all CSP constraints in \( C_V \) and all CSP constraints induced by the constraints on timelines in \( C_T \) and the assignment of \( V \) are satisfied.

It is important to note that the CNT framework is not included in the second kind of approach mentioned in the introduction. Indeed, we do not consider the planning problem over a fixed horizon, since dimension variables are actual variables on which constraints can be enforced and propagated. For example, consider a CNT containing one dimensional variable \( h \) of domain \( d(h) = [1..\infty] \), one timeline \( x = \{(0..2), h\} \), and three constraints \( x_1 = 0, x_2 = h, \) and \( vi \in \{1..h - 1\}, x_i + 1 - x_i \leq 1 \). Constraint propagation techniques can remove value 1 from the domain of \( h \), since if \( h = 1 \), then \( x_2 = 2 \) and \( x_1 = 0 \) are not compatible. As \( d(h) \) becomes \([2..\infty]\), constraint \( x_2 - x_1 \leq 1 \) must be satisfied. As \( x_1 = 0 \), constraint propagation can infer \( d(x_2) = \{0, 1\} \), which in turn allows value 2 to be removed from \( d(h) \). Therefore, constraints can be propagated in any direction and dimension variables will not necessarily be assigned first. Another useful feature of CNTs is the explicit presence of classical variables (outside timelines), which can model static features such as the choice of a download slot for a given observation. Last, CNTs can be easily extended to soft CNTs by replacing constraints by soft constraints, in order to model problems involving preferences such as the minimization of \( \text{card}\{o \in \{1..N_O\} \mid nd_o = 0\} \) if downloading all observations is not possible.

### Comparison with Existing Modeling

**Frameworks and Extensions**

The CNT framework is a kind of dynamic CSP, except that in CNTs, the number of potential variables may be unbounded, if the domain of a dimension variable is infinite. The interest of this “infinite” feature is that it makes it possible to model and solve, in a CSP-like way, planning problems over an initially unknown and unbounded horizon, or validation problems over an unbounded future. In dynamic CSPs, the number of potential variables and constraints is finite: variables are divided into a set of mandatory variables and a set of optional ones, and constraints are divided into a set of classical constraints and a set of activation constraints, which define when optional variables become active. Constraints are active only if their variables are active too. In CNTs, we do not explicitly define activation constraints: constraints are active depending on the domain of values of dimension variables. Another contribution of CNTs is that they explicitly identify the special role played by dimension variables in planning and scheduling problems. They allow global constraints which hold on a variable number of variables, such as \( \text{alldifferent}(x_i, i \in \{1..h\}) \), to be defined, whereas in dynamic CSPs, the scope of each constraint must be fixed before the resolution.

Compared to approaches completely integrating planning into constraint programming (third class given in the introduction), CNTs are built directly over variables and constraints, and not over more general entities such as intervals or structural constraints. This allows CNTs to be very generic, since any kind of constraint can be defined to model particular features of a real-world problem, such as global constraints or constraints involving both variables and t-variables. The generic aspect of
Important intermediate variables Given a timeline \( tl \) whose time reference is \( t \) and given a variable \( x \) whose domain is included in \( \mathbb{R} \), we can define intermediate variable \( \text{val}(tl, x) \) to represent the value of timeline \( tl \) at time \( x \). If \( tl \) has a piecewise constant evolution, then \( \text{val}(tl, x) = tl_i \) with \( i = \max(j \in [1..h(tl)] \mid t_j \leq x) \) if this quantity exists, \( \text{val}(tl, x) \) undefined otherwise. This definition is illustrated by Figure 1. Similarly, we can define intermediate variable \( \text{val}(b, x) \) to represent the value of timeline \( b \) just before time \( x \) (\( x \) excluded). If \( b \) has a piecewise constant evolution, then \( \text{val}(b, x) = b_i \) with \( i = \max(j \in [1..h(b)] \mid t_j < x) \) if this quantity exists, \( \text{val}(b, x) \) undefined otherwise. For timelines whose evolution is piecewise linear or discrete, \( \text{val}(tl, x) \) and \( \text{val}(b, x) \) are defined differently.

All these variables can be handled automatically in an efficient way, with dedicated global constraints on timelines hidden to the modeler, so that (s)he can directly use quantities \( \text{val}(tl, x) \) or \( \text{val}(b, x) \) to express constraints.

Other Modeling Examples

Before defining algorithms, let us show the modeling capabilities of CNTs on some problems from International Planning Competitions (IPCs). For each problem, different models can be defined. As in CSPs, finding a good model may not be straightforward. The models we present make some simplifications compared to the actual IPC formulations; in the experiments, we do not make such simplifications.

Domain Satellite (IPC3) A set of \( N_s \) satellites must take a set of \( N_t \) images. Each image \( im \) has a direction \( DI(im) \). With each satellite \( s \) is associated a set of observation instruments \( IN_s \), an instrument \( IN_i \), calibrated initially, and an initial pointing direction \( DI_s \). It is possible to compute a predicate \( SUPPORTS(im, in) \) which holds if image \( im \) can be performed with instrument \( in \). The duration needed to take a picture in a direction \( di \) with instrument \( in' \), starting from a direction \( di \) with instrument \( in \) calibrated, is denoted \( DU(di, in, di', in') \). It takes into account the necessity to calibrate \( in' \) if \( in' \neq in \). We denote by \( DU_{\text{min}} \) the minimum value of function \( DU() \).
To model this problem, we use one dimension variable \( h_s \) of domain \([0..N_s]\) per satellite \( s \). \( h_s \) represents the number of images taken by \( s \). Timelines \( im_s = \{(1..N_s), h_s\} \), \( in_s = (IN_s, h_s) \), \( dir_s = \{(DI(\im), im \in [1..N_s])\} \cup \{(DI_s), h_s\} \), and \( ts = (0..\{\max\}, h_s) \) represent respectively images taken by \( s \), associated instruments, associated directions, and the times when images are finished. Timelines \( in_s, dir_s, ts \) are initialized, so that variables \( in_{s,0}, dir_{s,0}, ts,0 \) can be used. A variable denoted \( tend \) represents the total duration needed to take all pictures. Different constraints are defined. Constraints \( \text{c9} \) and \( \text{c10} \) ensure that each image is taken exactly once, which prunes suboptimal solutions. Constraint \( \text{c11} \) defines the initial state. Constraint \( \text{c12} \) imposes feasibility constraints on the decisions. Constraints \( \text{c13} \) and \( \text{c14} \) describe the evolution of directions and times. Constraint \( \text{c16} \) is redundant but crucial for the algorithmic efficiency. Constraint \( \text{c15} \) defines \( tend \) as the makespan, which must be minimized.

\[
\begin{align*}
\sum_{s \in [1..N_s]} h_s &= N_t \quad (\text{c9}) \\
\text{alldifferent}(\{im_{s,i} \mid s \in [1..N_s], i \in [1..h_s]\}) & \quad (\text{c10}) \\
\forall s \in [1..N_s] ,
 dirs_0 &= \text{DI}(\text{IN}_s) \quad (\text{c11}) \\
\forall s \in [1..N_s], \forall i \in [1..h_s],
 \text{SUPPORT}(\text{init}_{s,i}, \text{in}_{s,i}) & \quad (\text{c12}) \\
\text{dir}_{s,i} &= \text{DI}(\text{in}_{s,i}) \quad (\text{c13}) \\
\text{ts}_{s,i} &= ts_{s,i-1} + DU(\text{dir}_{s,i-1}, \text{in}_{s,i-1}, \text{dir}_{s,i}, \text{in}_{s,i}) \quad (\text{c14}) \\
tend &\geq ts_{s,i} + (h_s - i) \cdot DU_{\text{min}} \quad (\text{c15}) \\
tend &= \max_{s \in [1..N_s]} ts_{s,h_s} \quad (\text{c16})
\end{align*}
\]

**Domain Trucks (IPC5)** The modeling of this domain shows how useful intermediate variables \( \text{val}(t, x) \) and \( \text{valb}(t, x) \) are. Domain Trucks involves a set of packages \( P \) and a set of trucks \( T \). Each truck \( \tau \in T \), initially located at a location \( L_I \), has a limited capacity and can load/unload packages, and drive between locations. Each package \( p \in P \) must be transferred from an initial location \( L_{I\tau} \) to a goal location \( L_{G\tau} \). Actions have durations and the goal is to minimize the makespan. In the sequel, we omit quantification on \( \tau \) and \( p \) in the expression of constraints.

For each truck \( \tau \), we use one dimension variable \( h_{\tau} \) representing the number of actions performed by \( \tau \) and a set of timelines \{\( \{\tau, a_{\tau}, p_{\tau}, t_{\tau}, n_{\tau}\}\) of dimension \( h_{\tau} \). For each \( i \in [1..h_{\tau}], a_{\tau,i} \) represents the action made by \( \tau \) at step \( i \) (load, unload, or drive), \( p_{\tau,i} \) is the package concerned by the action (if any), \( t_{\tau,i} \) is the start time of the action, and \( n_{\tau,i} \) are the location of \( \tau \) and the number of packages in \( \tau \) at the end of the action. Timeline \( t_{\tau} \) is the time reference of \( a_{\tau} \) and \( p_{\tau} \). For each package \( p \), we use one dimension variable \( h_p \) and two timelines \{\( \{t_{p,i}, l_{p,i}\}\) of dimension \( h_p \). For each step \( j \in [1..h_p], l_{p,j} \) represents the location of \( p \) at time \( t_{p,j} \). We consider that the location \( l_{p,j} \) can also be a truck. Timeline \( t_{p} \) is the time reference of \( l_{p} \). All timelines have a discrete evolution except from \( l_{p} \), whose evolution is considered to be piecewise constant. Timelines \( t_{\tau}, t_{p}, t_{\tau}, l_{p} \) are initialized.

Constraints \( \text{c20} \) to \( \text{c27} \) are imposed over these timelines. For example, \( \text{c21} \) defines the evolution of the number of packages in a truck. If a package is loaded by a truck, it must be at the same location as the truck just before the start of the loading \( \text{c22} \). If a package is unloaded by a truck, it is at the same location as the truck at the end of the unloading \( \text{c23} \). If a package is in a truck, then it has just been loaded and will be unloaded at the next step. Constraint \( \text{c24} \) is in this constraint, \( DU \) denotes the duration of an unloading. In fact, from the start of the loading to just before the end of the unloading, the package is considered to be in the truck. Constraint \( \text{c25} \) defines the makespan, and \( \text{c27} \) asserts both that a package must be at its goal location at the end and that when it is at its goal location, then its associated timeline is over. Other constraints are added to get a more efficient model. For example, \( \text{c25} \) prevents a package from being at the same place at two different steps, and \( \text{c29} \) is a transition constraint pruning suboptimal choices from the search space.

\[
\begin{align*}
(t_{\tau,0} &= t_{p,0} = 0) \land (l_{\tau,0} = L_{I\tau}) \land (l_{p,0} = L_{G\tau}) \quad (\text{c17}) \\
(t_{\tau,i} &\geq t_{\tau,i-1} + \text{duration}(a_{\tau,i}, t_{\tau,i-1}, t_{\tau,i}) \quad (\text{c18}) \\
(p_{\tau,i} = 0) &\leftrightarrow (a_{\tau,i} = \text{drive}) \quad (\text{c19}) \\
(l_{\tau,i} \neq l_{\tau,i-1} &\leftrightarrow (a_{\tau,i} = \text{drive}) \quad (\text{c20}) \\
\forall \tau_{i,i} &= n_{\tau_{i,i-1}} + f(t_{\tau_{i,i}}) \quad (\text{c21}) \\
&\text{with } f(\text{load}) = 1, f(\text{unload}) = -1, f(\text{drive}) = 0 \quad (\text{c22}) \\
(a_{\tau,i} = \text{load} &\leftrightarrow (\text{valb}(l_{\tau_{i,i}, t_{\tau_{i,i}}}, t_{\tau_{i,i}}) = t_{\tau_{i,i}-1}) \land (\text{val}(l_{\tau_{i,i}, t_{\tau_{i,i}}}, t_{\tau_{i,i}}) = \tau) \quad (\text{c23}) \\
(a_{\tau,i} = \text{unload} &\leftrightarrow (\text{val}(l_{\tau_{i,i}, t_{\tau_{i,i}}}, t_{\tau_{i,i}}) + DU = \tau) \land (\text{val}(l_{\tau_{i,i}, t_{\tau_{i,i}}}, t_{\tau_{i,i}}) = l_{\tau_{i,i}}) \quad (\text{c24}) \\
(l_{p,j} = \tau &\leftrightarrow (\text{val}(a_{\tau, t_{p,j}}, t_{p,j}) = l_{p,j}) \land (\text{val}(a_{\tau, t_{p,j}}, t_{p,j}) = p) \land (\text{val}(a_{\tau, t_{p,j}}, t_{p,j} + 1 - DU) = \text{unload}) \land (\text{val}(a_{\tau, t_{p,j} + 1 - DU} = p)) \quad (\text{c25}) \\
(l_{p,j} \in T &\leftrightarrow (l_{p,j+1} \notin T) \quad (\text{c26}) \\
\forall \text{temp}_{\in [1..N_p]} l_{p,\text{temp}} &\leftrightarrow (j = h_p) \quad (\text{c27}) \\
\text{alldifferent}(l_{p,i} \mid i \in [0..h_p]) &\quad (\text{c28}) \\
(a_{\tau,i} = \text{load} &\leftrightarrow (a_{\tau,i+1} \neq \text{unload}) \quad (\text{c29})
\end{align*}
\]
Function $\text{propagate}(V, C_V)$ transforms the CSP $(V, C_V)$ into an equivalent CSP $(V', C_V')$ by enforcing at least backward checking (Dechter 2003); this means that in $(V', C_V')$, all constraints whose scope is fully assigned are satisfied.

(R2) $\text{extend}(V', C_V, T, C_T, A_H)$ returns a pair $(V'', C_V'')$ such that CNTs $(V, C_V, T, C_T)$ and $(V', C_V', T, C_T)$ are equivalent, and such that, for every constraint $(S_V, S_T, fct) \in C_T$ for which there is a unique possible assignment $A$ for the dimension variables of timelines in $S_T$, $C_V''$ contains $fct(A)$.

To satisfy requirement (R1), $\text{propagate}$ can be any standard constraint propagation scheme, such as forward checking, arc consistency, or path consistency (Dechter 2003). Requirement (R2) can be fulfilled in different ways. The laziest version of $\text{extend}$ consists in generating constraints only when all dimension variables are assigned. The approach we use in the experiments is still lazy, but more incremental: when $\text{extend}(V', C_V', T, C_T, A_H)$ is called, it is possible to compare $A_H$, the previous minimum assignment of the dimension variables, and $A'_H$, the current minimum assignment of the dimension variables, and to add the set of t-variables $\{t_i | t_i \in T, i \in [A_H[h(t)] + 1..A'_H[h(t)])\}$ to $V'$. The way constraints are added to $C_V''$ depends on the type of constraint considered. For example,

- for a constraint such as $\forall i \in [1..h(t)], t_i \neq x$, function $\text{extend}$ can add the set of constraints $\{t_i \in [A_H[h(t)] + 1..A'_H[h(t)])\}$ to $C_V''$;
- a constraint like $\text{allDifferent}(t_i | i \in [1..h(t)])$ can generate constraint $\text{allDifferent}(t_i | i \in [1..A'_H[h(t)])$ if $A_H[h(t)] \neq A'_H[h(t)]$.

In fact, constraints can be added as soon as they must necessarily be satisfied. The design of specialized schemes for function $\text{extend}$ for a constraint on timelines $(S_V, S_T, fct)$ can be highly dependent on $fct$ and is not discussed here.

Functions $\text{dynDFS}$ and $\text{recDynDFS}$ Given a CNT $(V, C_V, T, C_T)$, the systematic depth-first tree search is performed by calling $\text{dynDFS}(V, C_V, T, C_T)$. After an initial extension/propagation step (lines 4 and 5), function $\text{dynDFS}$ calls function $\text{recDynDFS}$ if the initial problem has not been proved to be inconsistent, and returns null otherwise.

If it terminates, $\text{recDynDFS}(V, C_V, T, C_T)$ returns an optimal consistent assignment $A$ of the CNT $(V, C_V, T, C_T)$ if it exists, and null otherwise. If there is a unique possible assignment of $V$, this assignment is returned (lines 12-13). Otherwise, the algorithm chooses a variable $x \in V$ not assigned yet and builds a partition of the domain of $x$, according to some heuristics (lines 15-16). The two search subspaces defined by this partition are then successively explored (lines 18 to 23). For each of them, $\text{recDynDFS}$ first propagates constraints using $\text{dynPropagate}$ (line 21). If no inconsistency is revealed (line 22), $\text{dynDFS}$ is recursively called (line 23). If a solution $A' \neq \text{null}$ is returned, it is recorded as well as the best value known for the objective.

Discussion and properties Algorithm $\text{dynDFS}$ is a generic algorithm which covers several existing approaches. Indeed, approaches reasoning over a sequence of size-bounded CSPs simply correspond to variable/value choice heuristics (lines 15 and 16) where all dimension variables are assigned first, with their minimal values. $\text{dynDFS}$ can
also adopt a strategy where horizons are dynamically incremented during search, when constraint propagation prunes the minimum value in the domain of dimension variables. As a result, dynDFS allows several approaches to be compared inside a common framework.

Formal properties of dynDFS are given below. This algorithm is correct but does not necessarily terminate, since it might get trapped in infinite branches of the search space when the domain of some variable is infinite.

**Proposition 1 (Correctness)** If functions propagate and extend satisfy (R1) and (R2), then dynDFS is correct: if it terminates, its result is an optimal consistent assignment if the CNT considered admits a solution and null otherwise.

**Proposition 2 (Termination)** If all domains of values are finite, then dynDFS terminates. If all non-dimension variables have a finite domain and if the problem admits at least one solution, then there exist choice heuristics (lines 13 and 16) such that dynDFS finds a consistent assignment in a finite time. In general, dynDFS does not terminate.

**Proposition 3 (Complexity class of CNTs)** Deciding whether there exists a CNT assignment with an objective value lesser than a given threshold 0 is (a) NP-complete for CNTs where all domains of values are finite, (b) semi-decidable for CNTs such that all non-dimension variables have a finite domain, and (c) undecidable in general.

Sketch of the proofs: for Prop. 1 the idea is to prove that if it terminates, dynPropagate($V, C_V, T, C_T$) returns a couple ($V', C_V'$) such that ($V, C_V, T, C_T$) and ($V', C_V'$) are equivalent, and that if it terminates, recDynDFS($V, C_V, T, C_T$) returns an optimal consistent assignment of ($V, C_V, T, C_T$) if there exists one and null otherwise. For Prop. 2 if all domains of non-dimension variables are finite, it suffices to use an assignment heuristics that iteratively increments the maximum value that can be assigned to a dimension variable. For Prop. 3 checking the consistency of a CNT assignment is polynomial and any finite CSP can be expressed as a CNT; hence the NP-completeness result; if all non-dimension variables have a finite domain, then Prop. 2 implies the semi-decidability result; for undecidability in general, it was shown in [Verfaille, Pralat, & Lemaître 2008] that the halting problem can be expressed as the problem of finding a consistent assignment of a CNT.

**Experiments**

To measure the practical interest of CNTs, we performed experiments on domains BlocksWorld (IPC2), Satellite (propositional and simpletime versions, IPC3), and Trucks (propositional and temporal versions, IPC5). The first task was to build CNT representations manually as described previously. For these domains, the value of dimension variables can be bounded while preserving optimality, hence dynDFS terminates. The goal is to minimize the makespan.

The ideas of CNT and dynDFS are implemented over Choco (Laburthe 2000), a constraint programming library. The constraint propagation algorithm used is GAC (Generalized Arc Consistency [Dechter 2003]). Function extend is implemented via constraints ifThen($h \geq i, c_i$), which activate constraint $c_i$ only when guard $h \geq i$ holds. Other more efficient implementations of function extend could be developed. Several parameter settings were tested for the choice of the variable to consider at each step: (1) consider dimension variables first; (2) consider non-dimension variables first; (3) consider dimension and non-dimension variables in any order. The results presented for Trucks correspond to option (2), and by considering first variables having a minimum domain size for non-dimension variables, and variables having a minimum minimal value for dimension variables.

We ran our experiments on an AMD Opteron processor, 2.4 GHz, with 1GB RAM, under Linux, with a time limit of half an hour per problem. We compared dynDFS(CNT) with the optimal planners awarded at the last planning competition: MaxPlan, SatPlan, and CPT. MaxPlan and SatPlan can handle propositional domains. CPT can handle both propositional and temporal domains. Table 1 shows that in general, dynDFS(CNT) performs better than MaxPlan, SatPlan, and CPT. On small-size instances, dynDFS(CNT) can be slower since, as it contains more information, the initialization can be longer. On harder instances, dynDFS(CNT) provides significant gains. Instances of BlocksWorld are easy for dynDFS(CNT) thanks to symmetry breaking constraints and to constraints forcing necessary moves to be done. Instances of Satellite, propositional or temporal, are solved in a few seconds with dynDFS(CNT), whereas with MaxPlan, SatPlan, and CPT, which work on models containing less information, they are solved only in several minutes or unsolved at all. Trucks appears to be more challenging, in the sense that the CNT representation speeds search, but does not modify the intrinsic complexity of the problem. For unsolved instances, as shown in Figure 2 dynDFS(CNT) is able to quickly produce solutions whose quality is better than the quality of the solution produced by SGPlan, a heuristic-based planner awarded in IPC5. In fact, for domain Trucks, dynDFS(CNT) is quite anytime: the optimal solution is reached quite quickly, and the rest of the time is dedicated to prove optimality.

**Conclusion**

In this paper, we presented Constraint Networks on Timelines (CNTs), a generic constraint-based framework for modeling and solving planning and scheduling problems. This framework is compact and has a clear semantics based on variables and constraints. A generic dynamic depth-first tree search algorithm using constraint propagation has been developed and tested on several instances taken from planning competitions. Experimental results have shown the practical interest of the approach, both compared to existing

---

For MaxPlan, see http://www.cse.wustl.edu/~chen/maxplan/. For SatPlan, see http://www.cs.rochester.edu/~kautz/satplan. For CPT, see http://www.cril.univ-artois.fr/~vidal/#cpt. For CPT, we use CPT1 because CPT2 is not publicly available.
optimal planners in terms of time to get the optimal solution and to prove optimality, and compared to heuristic planners in terms of solution quality. In particular, some problems unsolved by existing optimal planners are solved in a few seconds with CNTs. In the end, the basic constraint-based semantics allows various kinds of information to be captured in CNTs, such as constraints modeling scheduling aspects as well as planning aspects, temporal constraints, constraints on both dimension and timeline variables, or constraints on binary or n-ary variables. The key factor explaining the algorithmic success of CNTs is that they allow efficient models containing information such as global constraints, constraints between states, constraints between actions, symmetry breaking constraints, constraints pruning suboptimal solutions, or redundant constraints, to be developed. Exploiting the information available avoids the planner from being blind, while preserving optimality.

In the future, we believe that the performance of algorithms on CNTs could be improved significantly, since constraint programming techniques such as intelligent backtracking, structural decomposition, improved heuristics, limited discrepancy search, soft constraint propagation, constraint preprocessing, or randomization and restart, have not been used yet. It would also be interesting to develop approximate algorithms and to compare their performance with heuristic-based planners. Last, the approach should be extended in order to be able to handle uncertainty.

![Figure 2: Evolution of the makespan given by dynDFS(CNT) and comparison with the makespan given by an heuristic-based planner (SGPlan, which uses FF), on problems trucks05-temporal (left) and trucks06-temporal (right).](image)

**Table 1: Comparison between dynDFS(CNT) and some optimal planners, on propositional and temporal domains.**

<table>
<thead>
<tr>
<th>Problem</th>
<th>CPU time (sec.)</th>
<th>MaxPlan</th>
<th>SatPlan</th>
<th>CPT</th>
<th>dynDFS(CNT)</th>
<th>Makespan</th>
</tr>
</thead>
<tbody>
<tr>
<td>bw-large-a</td>
<td>0.51</td>
<td>0.38</td>
<td>0.14</td>
<td>1.08</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>bw-large-b</td>
<td>4.64</td>
<td>2.36</td>
<td>0.96</td>
<td>2.60</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>bw-large-c</td>
<td>171.19</td>
<td>38.99</td>
<td>56.17</td>
<td>6.91</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>bw-large-d</td>
<td>-</td>
<td>455.65</td>
<td>-</td>
<td>15.34</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>bw-ipc10</td>
<td>0.47</td>
<td>0.24</td>
<td>0.03</td>
<td>0.79</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>bw-ipc20</td>
<td>-</td>
<td>5.42</td>
<td>407.20</td>
<td>1.48</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>bw-ipc30</td>
<td>-</td>
<td>44.28</td>
<td>-</td>
<td>2.55</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>bw-ipc40</td>
<td>-</td>
<td>183.91</td>
<td>151.33</td>
<td>4.92</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>bw-ipc50</td>
<td>-</td>
<td>-</td>
<td>8.39</td>
<td>-</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>satellite05-prop</td>
<td>0.56</td>
<td>0.41</td>
<td>0.38</td>
<td>0.52</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>satellite06-prop</td>
<td>0.32</td>
<td>0.47</td>
<td>0.16</td>
<td>0.54</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>satellite07-prop</td>
<td>0.46</td>
<td>0.62</td>
<td>0.20</td>
<td>0.58</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>satellite08-prop</td>
<td>25.45</td>
<td>40.73</td>
<td>92.23</td>
<td>1.53</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>satellite09-prop</td>
<td>1.76</td>
<td>2.77</td>
<td>1.61</td>
<td>0.99</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>satellite10-prop</td>
<td>54.28</td>
<td>16.58</td>
<td>40.83</td>
<td>0.78</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>satellite11-prop</td>
<td>7.87</td>
<td>12.28</td>
<td>4.40</td>
<td>0.78</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>satellite12-prop</td>
<td>556.40</td>
<td>111.72</td>
<td>-</td>
<td>8.35</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>satellite13-prop</td>
<td>644.82</td>
<td>308.38</td>
<td>-</td>
<td>0.70</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>satellite14-prop</td>
<td>95.80</td>
<td>74.37</td>
<td>-</td>
<td>8.82</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>truck01-prop</td>
<td>0.44</td>
<td>2.26</td>
<td>-</td>
<td>4.37</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>truck02-prop</td>
<td>39.95</td>
<td>60.74</td>
<td>-</td>
<td>22.66</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>truck03-prop</td>
<td>1163.32</td>
<td>409.21</td>
<td>-</td>
<td>7.57</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>truck04-prop</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>45.97</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>truck05-prop</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>463.74</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>truck06-prop</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1780.72</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>truck07-prop</td>
<td>709.74</td>
<td>642.51</td>
<td>-</td>
<td>8.42</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

References


