

# Constraint Bag Process Model: An Interdisciplinary Process Mining Approach to Lean Construction

Daniel Pérez

Department of Construction Engineering  
and Management  
School of Engineering  
Pontificia Universidad Católica de Chile  
Santiago, Chile  
duperez@uc.cl

Camilo Ruiz-Tagle

Department of Computer Science  
School of Engineering  
Pontificia Universidad Católica de Chile  
Santiago, Chile  
cruiztagle@uc.cl

Jorge Muñoz-Gama

Department of Computer Science  
School of Engineering  
Pontificia Universidad Católica de Chile  
Santiago, Chile  
jmun@uc.cl

Michael Arias

Department of Business Computer Science  
Universidad de Costa Rica,  
Sede de Occidente  
Costa Rica  
michael.arias\_c@ucr.ac.cr

Luis Fernando Alarcón

Department of Construction Engineering  
and Management  
School of Engineering  
Pontificia Universidad Católica de Chile  
Santiago, Chile  
lalarcon@ing.puc.cl

Marcos Sepúlveda

Department of Computer Science  
School of Engineering  
Pontificia Universidad Católica de Chile  
Santiago, Chile  
marcos@ing.puc.cl

**Abstract**—Computer science techniques, methodologies, and approaches, are directly applied to improve other engineering disciplines. Construction is no exception, where software and data analysis are used to improve the processes of control and monitoring of construction projects. Most of the analysis are based on a key-value perspective analysis of the data. However, an emerging Process Mining discipline has proven to be able to capture a different process perspective. This interdisciplinary work is a step on this direction, proposing the Constraint Bag Process Model (CBPM), a novel use of Process Mining for answering process-oriented questions on construction projects.

**Index Terms**—Process Mining, Lean Construction, Last Planner System, Interdisciplinary

## I. INTRODUCCIÓN

La industria de la construcción en Chile representa cerca del 6,4 % de la economía nacional y su alza tiene estrecha relación con el crecimiento y calidad de vida del país [1]. El control y administración de su producción resulta ser de suma importancia en el proceso constructivo con el fin de aumentar la productividad en la industria. Así, es posible disminuir riesgos y costos al consumidor y la sociedad [2]. Dentro de los múltiples mecanismos de control y planificación es posible encontrar *Last Planner System (LPS)*, basado en la filosofía *Lean Construction* (i.e., Construcción Sin Pérdidas, práctica que busca minimizar pérdidas en materiales, tiempo y esfuerzo aplicado con el fin de aumentar la producción y generación de valor [3]). Esta metodología ha permitido mejorar índices de producción en hasta un 30 %, mejorando retornos económicos y estabilizando flujos de trabajo [4].

El uso de técnicas computacionales para la mejora en los procesos de control y seguimiento de proyectos de construcción se ha centrado principalmente en un análisis orientado a datos con una perspectiva clave-valor (e.g., ¿Qué número de compromisos se han incumplido en la planificación inicial de esta semana?, ¿Qué equipo se demora más en el cumplimiento de plazos?). Un buen ejemplo de esta perspectiva de análisis es el uso de técnicas de *Big Data* para la toma de decisiones en proyectos con gran cantidad de datos disponibles [5]. El auge del uso de sistemas de información ha permitido que disciplinas como la ciencia orientada a los datos (*Data Science*) lideren un enfoque más orientado a los datos en busca de obtener valor a partir del análisis de esa gran cantidad de datos disponibles [6]. Sin embargo, existe una perspectiva orientada a procesos [7] que no ha sido tan considerada y sistematizada con el objetivo de mejorar los procesos desde que inician hasta que terminan. La detección y análisis de patrones de procesos (e.g., la trayectoria seguida de cumplimiento, o incumplimiento, de los compromisos adquiridos inicialmente a lo largo de un mes de proyecto), pueden enriquecer y proporcionar un valor agregado a los análisis actuales enfocados en los procesos. En los últimos años, *Process Mining* [6], [8] ha emergido como una disciplina dentro de la ciencia orientada a los procesos (*Process Science*) y la computación capaz de analizar los datos desde esta perspectiva de procesos [9]. En concreto, *Process Mining* incluye técnicas y algoritmos para *descubrir* de forma automática modelos de proceso a partir de los datos, *verificar la conformidad* entre un modelo de procesos dado y los datos, o *extender* un modelo dado con información detectada en los datos [6].

El presente artículo propone el uso de *Process Mining* para

el análisis de la perspectiva de procesos en proyectos de construcción. Sin ir más lejos, el artículo define un nuevo modelo llamado *Constraint Bag Process Model (CBPM)*, el cual permite realizar de forma intuitiva un análisis de las trayectorias de cumplimiento de un conjunto de compromisos (i.e., restricciones) planificadas en proyectos de construcción, condicionadas a atributos como responsable, prioridad, tipo de restricción, entre otros. Finalmente, el modelo permite detectar trayectorias específicas dentro de una implementación de metodología *Last Planner System (LPS)*.

El resto del artículo se estructura de la siguiente manera: la Sección II presenta una breve introducción a los conceptos de *Last Planner Systems (LPS)*. Luego, la Sección III presenta la conceptualización del modelo *Constraint Bag Process Model*, y la Sección IV muestra su aplicación en un contexto concreto de LPS llamado IMPERA. Finalmente, se presenta el trabajo relacionado en la Sección V, y las conclusiones y el trabajo futuro en la Sección VI.

## II. LAST PLANNER SYSTEM (LPS)

*Last Planner System (LPS)*, o Sistema del Último Planificador, es una metodología basada en la gestión de compromisos y asignaciones desarrollada en 1994 por Glenn Ballard. Desde su creación, ha sido utilizada en más de veinte países del mundo [10] obteniendo importantes mejoras a nivel de producción y retornos económicos en la industria de la construcción. Su estructura está basada en cuatro fases principales, tal como se muestra en la Figura 1.



Figura 1. Etapas de LPS

Las etapas ilustradas en la Figura 1 corresponden a [11]:

- *Plan Maestro*: tiene como propósito establecer los principales hitos del programa a realizar. Se detallan las fechas más importantes del proyecto.
- *Planificación de Fases*: contempla un mayor nivel de detalle. Se especifican fechas de traspaso de actividades. Su realización tiene un fuerte énfasis en el trabajo colaborativo.
- *Lookahead*: una de las principales etapas de la metodología. Consiste en, de forma colaborativa, encontrar restricciones y factibilidad para las tareas a realizar en un periodo  $t$  de tiempo futuro (i.e, típicamente un *Lookahead* comprende un periodo entre 4-6 semanas). En esta etapa, se establecen compromisos y fechas de liberación de restricciones. Entenderemos por restricción a *un elemento o requisito que evita el comienzo, avance o término de una actividad según lo planeado. Una vez que esta esté liberada, significa que la actividad puede comenzar, continuar o terminar correctamente.*
- *Programación Semanal*: concentra el mayor nivel de detalle de planificación. Establece el Inventario de Trabajo Ejecutable (ITE) para la semana. Además, se establecen

indicadores que permitan conocer el estado actual del programa y/o proyecto.

En concreto, el modelo propuesto en este artículo se centra en la etapa *lookahead*. El modelo permitirá analizar de forma intuitiva las trayectorias de cumplimiento de las restricciones dentro de los periodos de *lookahead*.

Por simplicidad el modelo presentado en este artículo se basa en la metodología LPS. Sin embargo, los conceptos definidos en este trabajo pueden ser fácilmente adaptados y aplicados a otras metodologías *Lean Construction* basadas en restricciones, e incluido, en otros contextos.

## III. CONSTRAINT BAG PROCESS MODEL (CBPM)

En esta sección se define el modelo *Constraint Bag Process Model (CBPM)*. Este modelo se inspira en el concepto de una *mochila de restricciones*, i.e., una carga de trabajo que, de no cumplirse, se va acumulando en un actor determinado del proyecto.

Nótese, que el concepto de definir un modelo de proceso es un abuso del lenguaje. En concreto, esta sección propone una instanciación específica de la estructura de datos requerida por los algoritmos de *Process Mining* para el contexto de proyectos de construcción [6]. En concreto, los elementos necesarios a definir son: 1) el concepto de **caso** o instancia del proceso (Sección A), 2) la definición de **actividades** del proceso (Sección B), y 3) el **orden** relativo entre actividades (Sección C). Adicionalmente, tanto los casos como las actividades son enriquecidas con **atributos** adicionales para su análisis (Sección D).

Según la definición CBPM, el resultado de aplicar algoritmos de descubrimiento de procesos sobre los datos estructurados, da como resultado el modelo del proceso a ser analizado. Dado el carácter interdisciplinar del método, y el sesgo representativo del modelo CBPM (e.g., por definición no puede existir paralelismo entre actividades), la recomendación es aplicar algoritmos con bajo nivel de complejidad, y con una salida fácilmente interpretable. En ese caso, buenos ejemplos de algoritmos a utilizar son Heuristic Miner [12], Inductive Miner [13], o alternativas comerciales similares como las proporcionadas por Disco [14] y Celonis [15], [16].

### A. Definición de Caso

Entenderemos como caso a la ejecución de un conjunto de restricciones por una persona (o responsable, según LPS) en el periodo de tiempo comprendido en un *Lookahead*. A su vez, cada *Lookahead* comprende un periodo entre 4 y 6 semanas, dependiendo de la complejidad y decisión propia de cada proyecto en particular.

En otras palabras, un identificador de caso se define como una tupla  $(l, r)$  donde  $l$  indica el *lookahead* concreto del caso, y  $r$  corresponde al *id* del responsable del caso. Formalmente, definimos el conjunto de identificadores de caso  $C$  como:

$$C = \{(l, r) \mid l \in L \wedge r \in R\}$$

donde  $L$  y  $R$  corresponde al conjunto de *lookheads* y responsables considerados, respectivamente.

## B. Definición de Actividades

Luego de definir los casos, es necesario definir las actividades. Las actividades permitirán representar la criticidad de la semana en función de los estados actuales de las restricciones, y de la importancia con que se encuentren tipificadas (ALTA, MEDIA, BAJA).

Con los estados actuales de las restricciones, semana a semana, y su relación con su importancia, es posible definir un *color* que representará el estado de la *mochila* de restricciones para cada caso. Por lo tanto, las actividades o estados para cada caso son: *GREEN*, *YELLOW* y *RED*. Donde *GREEN* es el estado de menor criticidad, *YELLOW* el de mediana criticidad y *RED* el de mayor criticidad. Estas actividades son calculadas en base al estado e importancia de las restricciones en la semana  $i$  de cada *Lookahead*.

**B1. Estado de las Restricciones:** Los estados de las restricciones vienen definidos según la metodología LPS. Estos son:

- Pendiente en Plazo (PP): una restricción está *Pendiente en Plazo* cuando no ha sido liberada y aun no cumple su plazo máximo de liberación comprometido.
- Atrasada (ATR): una restricción está *Atrasada* cuando el plazo máximo de liberación comprometido se ha cumplido y la restricción aún no ha sido liberada.
- Liberada a Tiempo (LAT): cuando la restricción fue liberada antes de cumplirse el plazo máximo de liberación comprometido.
- Liberada con Atraso (LATR): cuando una restricción ha sido liberada después de haberse cumplido el plazo máximo de liberación comprometido.

De este modo, cuando se instancia una restricción, dado un responsable, en el contexto de la primera semana, son creadas en el estado inicial *Pendiente en Plazo (PP)*. Después de esto, cuando avanzamos a la semana siguiente, las restricciones pueden tomar el valor de *liberada a tiempo (LAT)*, *liberada con atraso (LATR)* o *atrasada (ATR)*, según el hito temporal de entrega comprometido. En este trabajo, el estado de las restricciones es importante para entender las reglas con que definimos la relevancia de éstas.

**B2. Importancia de las Restricciones:** Las restricciones que conforman un caso poseen distintos *niveles de importancia (ALTA, MEDIO, BAJA)*. Dicho nivel, junto con el estado de las restricciones, condiciona la forma como se categorizan las actividades (*GREEN, YELLOW, RED*).

En nuestro modelo nos hemos enfocado en aquellas restricciones que aún no han sido liberadas; *pendiente en plazo (PP)* y *atrasadas (ATR)*, y para ellas definimos los siguientes criterios ilustrados en las Tablas I y II.

La lectura de las Tablas I y II es la siguiente: para el caso de porcentaje de restricciones en *ALTA*, entre 20% y 50%, y si el complemento distribuye en –a lo menos– un 80% de *MEDIAS* y un 20% de *BAJAS*, entonces el estado es *RED*. En otras combinaciones de restricciones *MEDIAS* y *BAJAS*, el estado será *YELLOW*.

Tabla I  
REGLAS PARA RESTRICCIONES PENDIENTES EN PLAZO.

ALTA	MEDIO	BAJO	COLOR
>50 %	(0 %-50 %)	0 % otros casos	RED YELLOW
20-50 %	>80 %*(100 %- %Altas)	<20 %*(100 %- %Altas) otros casos	RED YELLOW
20 %	>80 %*(100 %- %Altas)	<20 %*(100 %- %Altas) otros casos	YELLOW GREEN

Tabla II  
REGLAS PARA RESTRICCIONES ATRASADAS.

ALTA	MEDIO	BAJO	COLOR
>50 %	(0 %-50 %)	0 % otros casos	RED YELLOW
20-50 %	>80 %*(100 %- %Altas)	<20 %*(100 %- %Altas) otros casos	RED YELLOW
20 %	>80 %*(100 %- %Altas)	<20 %*(100 %- %Altas) otros casos	YELLOW GREEN

## C. Orden de actividades

Cada una de las actividades del proceso representa una de las semanas dentro del *lookahead*. Consecuentemente, se utiliza el orden entre las semanas como el orden entre actividades. Formalmente, definimos una instancia del proceso  $\sigma$  como la secuencia ordenada:

$$\sigma = \langle a_1, a_2, \dots, a_n \rangle \text{ donde } a_i \in A$$

donde  $n$  es el número de semanas del *lookahead* (e.g., 4 o 5), y  $a_i$  representa la actividad de la semana  $i$  dentro del universo  $A$  de actividades definidas (i.e., *GREEN, YELLOW, RED*).

## D. Atributos de Caso y Actividad

Tanto las actividades como los casos se encuentran enriquecidos en su caracterización mediante atributos adicionales que representan características más específicas del negocio.

**D1. Caso:** incluye atributos como *id* de caso, *id* de responsable y su nombre correspondiente.

**D2. Actividades:** se agregan a los atributos adicionales el *id* de la semana dentro del *lookahead*, fecha de inicio y término del mismo, caracterización de las restricciones según importancia (*ALTA, MEDIA, BAJA*) y según el tipo (*Diseño, Condiciones Previas, Mano de Obra, Materiales, Otros*).

## IV. EL CASO IMPERA

La presente sección muestra el despliegue del modelo CBPM aplicando la metodología LPS, con datos obtenidos a partir de un software de gestión y planificación llamado IMPERA. A partir de su aplicación, se realiza un experimento y se analizan diferentes tipos de preguntas a responder usando el modelo propuesto.

### A. Contexto

IMPERA es un software desarrollado para la implementación de LPS creado en 2002 por el Centro de Excelencia en Gestión de Producción de la Pontificia Universidad Católica de Chile (GEPUC). La herramienta ha sido utilizada en diversos países de Latinoamérica [17], creando más de 80 indicadores

que apoyan a la toma de decisiones del proyecto. En el presente informe, para el modelo de CBPM, se utilizan diversos proyectos correctamente documentados en la plataforma de IMPERA.

### B. Datos

En el contexto del experimento a realizar, se consideró la generación de datos correspondientes a 200 casos para diez responsables en 20 *lookaheads*. En este caso, la duración de cada *lookahead* se ha fijado en 4 semanas. Se generaron los comportamientos de las variables consideradas teniendo especial cuidado en que su comportamiento sea representativo a los datos de IMPERA.

### C. Método

Posterior a la obtención de los datos para este experimento, se utiliza el algoritmo de descubrimiento (conocido como *process graph*) de la suite *Celonis Process Mining*<sup>1</sup>. Este es un algoritmo de la familia *Heuristic Miner* [6] y su '*representative bias*' permite obtener resultados claros y entendibles del flujo del proceso para mejorar la toma de decisiones. Además, permite la configuración de distintos *dashboards* personalizables para analizar más a fondo los datos dependiendo de las necesidades del usuario.

### D. Resultados y Discusión

El algoritmo de descubrimiento de procesos de *Celonis* [15], [16] permitió obtener un modelo *CBPM* (cf. Figura 2), representación de los diversos patrones de comportamiento con que se relacionan las actividades definidas y las transiciones entre actividades.

A partir de este modelo, se pueden definir preguntas desde la perspectiva de procesos con interés para los responsables del proyecto. En concreto, se pueden definir preguntas sobre los casos y sus características con 1) presencia de alguna actividad, 2) presencia de una transición concreta entre actividades, y 3) presencia de patrones de transición más complejos.

En concreto, analizando el contexto de proyectos de construcción, se proponen las siguientes preguntas a modo ilustrativo:

- ¿Cómo se comportan las trazas (*lookahead*) que pasan por una Actividad *RED*?
- ¿Cómo se comportan los casos en que hay un cambio de *GREEN* a *RED*? ¿Logran recuperarse?
- ¿Cómo se comportan los casos en que hay loops entre estados? ¿Cuáles se "recuperan"? ¿Cuáles empeoran?

Nótese que, dada las limitaciones del papel, el análisis de los resultados que se puede hacer sobre él se hace de manera estática, listando las variantes y mostrando los resultados filtrados. Sin embargo, el análisis propuesto en el modelo *CBPM* se realizaría de una forma interactiva, estableciendo filtros en tiempo real (*on-the-fly*) sobre actividades o transiciones, o bien, filtros utilizando criterios más complejos (e.g., actividades con patrones como *eventually follow*).

<sup>1</sup>Celonis tool: <http://www.celonis.com/en/product>

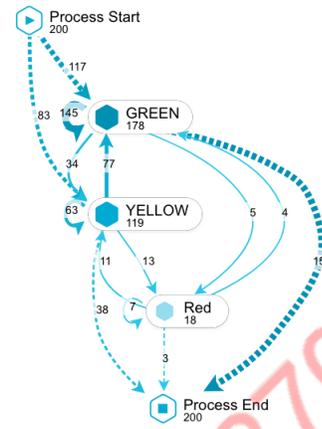


Figura 2. Proceso descubierto para todas las variantes.

D1. ¿Cómo se comportan las trazas (*lookahead*) que pasan por una Actividad *RED*?: Como se puede ver en la Figura 3, todos los casos que pasan por la Actividad *RED*. Dentro de estos casos, se pueden identificar 3 grupos: Casos que terminan en *YELLOW*; Casos que terminan en *RED*; y Casos que terminan en *GREEN*. Nos centraremos en los casos que terminan en *RED* y en *YELLOW*.

Para los casos que terminan en *YELLOW*, las restricciones del tipo *Condiciones previas* son dominantes. Estas restricciones poseen importancia *ALTA* como se puede apreciar en la Tablas III, IV y la Figura 4.

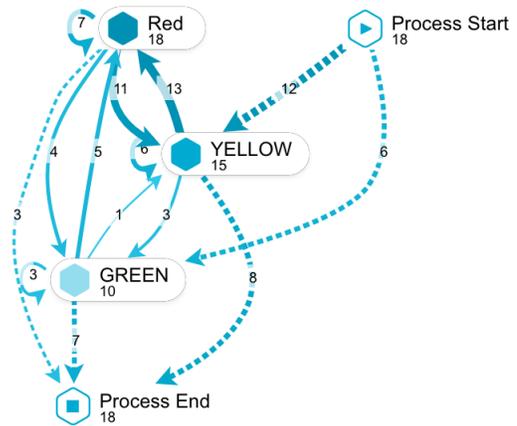


Figura 3. Proceso descubierto para la pregunta 1

Para los casos que terminan en *RED*, las restricciones del tipo *MATERIALES* y *M.O* son dominantes. Estas restricciones poseen importancia *ALTA* como se puede apreciar en la Tablas V, VI y la Figura 5. Analizando todas estas trazas que pasan por *RED*, es posible apreciar que en el 38% de los casos terminan de manera óptima (en el estado *GREEN*). El restante termina con un estado *YELLOW* o *RED*. Para el primer caso, es posible apreciar que el 75% de estos vuelven a pasar una segunda vez por el estado *RED*.

Tabla III  
CASOS POR TIPO DE RESTRICCIONES QUE FLUYEN A TRAVÉS DE RED Y TERMINAN EN YELLOW

Variante	%C.Previas	%M.O.	%Materiales	%Diseño	%Otros
G-R-Y-Y	50 %	50 %	0 %	0 %	0 %
G-R-R-Y	57 %	43 %	0 %	0 %	0 %
Y-R-Y-Y	67 %	33 %	0 %	0 %	0 %
G-Y-R-Y	71 %	29 %	0 %	0 %	0 %
Y-R-R-Y	83 %	0 %	17 %	0 %	0 %

Tabla IV  
CASOS POR IMPORTANCIA DE RESTRICCIONES QUE FLUYEN A TRAVÉS DE RED Y TERMINAN EN YELLOW

Variante	%ALTA	%MEDIAS	%BAJAS
G-Y-R-Y	%43	%0	%57
Y-R-Y-Y	%75	%25	%0
G-R-Y-Y	%75	%0	%25
Y-R-R-Y	%83	%17	%0
G-R-R-Y	%86	%0	%14

Tabla V  
CASOS POR TIPO DE RESTRICCIONES QUE FLUYEN A TRAVÉS DE RED Y TERMINAN EN RED

Variante	%C.Previas	%M.O.	%Materiales	%Diseño	%Otros
Y-R-R-R	0 %	43 %	43 %	0 %	14 %
Y-Y-R-R	0 %	0 %	100 %	0 %	0 %

Tabla VI  
CASOS POR IMPORTANCIA DE RESTRICCIONES QUE FLUYEN A TRAVÉS DE RED Y TERMINAN EN RED

Variante	%ALTA	%MEDIAS	%BAJAS
Y-Y-R-R	%100	%0	%0
Y-R-R-R	%100	%0	%0

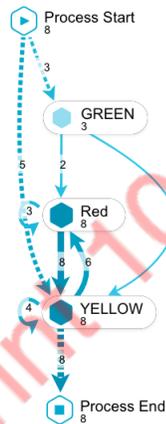


Figura 4. Proceso descubierto para los casos que fluyen por RED y terminan en YELLOW.

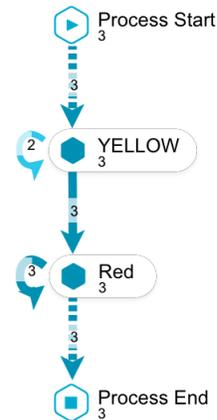


Figura 5. Proceso descubierto para los casos que fluyen por RED y terminan en RED.

D2. ¿Cómo se comportan los casos en que hay un cambio de GREEN a RED? ¿Logran recuperarse?: Estos casos son representados en las Tablas VII, VIII y en la Figura 6. Como se puede apreciar, la tendencia es que las restricciones del tipo COND.PREVIAS sea dominante, junto con las restricciones de importancia ALTA.

D3. ¿Cómo se comportan los casos en que hay loops entre estados? ¿Cuáles se “recuperan”? ¿Cuáles empeoran?:

Se pueden descubrir los siguientes Loops en la figura 2: GREEN- YELLOW, YELLOW-GREEN; GREEN- RED, RED-GREEN ; GREEN- GREEN ;YELLOW- YELLOW ;RED-RED. Centramos nuestro análisis en los loops GREEN- RED, RED-GREEN y RED- RED.

Respecto de los loops GREEN- RED, RED-GREEN , son casos que terminan en GREEN, poseen un 80% de Condiciones Previas. Un 60% de ellas tienen importancia ALTA.

Tabla VII  
CASOS POR TIPO DE RESTRICCIONES QUE TIENEN SECUENCIA GREEN-RED

Variante	%C.Previas	%M.O.	%Materiales	%Diseño	%Otros
G-R-R-Y	57 %	43 %	0 %	0 %	0 %
G-R-G-G	80 %	0 %	20 %	0 %	0 %
G-R-Y-Y	50 %	50 %	0 %	0 %	0 %

Tabla VIII  
CASOS POR IMPORTANCIA DE RESTRICCIONES QUE TIENEN SECUENCIA GREEN-RED

Variante	%ALTA	%MEDIAS	%BAJAS
G-R-G-G	%60	%0	%40
G-R-Y-Y	%75	%0	%25
G-R-R-Y	%86	%0	%14

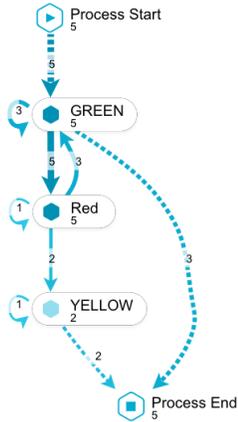


Figura 6. Proceso descubierto para los casos con cambio GREEN a RED.

Como este caso corresponde a una sola variante (G-R-G-R), su tendencia se puede apreciar en la Figura 7.

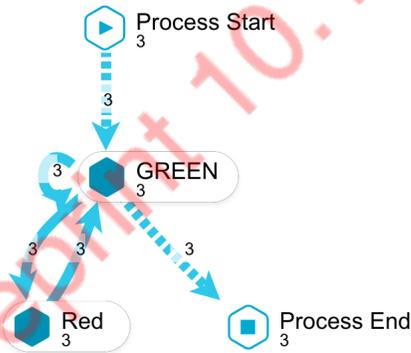


Figura 7. Proceso descubierto para los casos con loops GREEN-RED, RED-GREEN.

Respecto de los loops RED- RED (ver Figura 8), se puede apreciar que las restricciones de importancia ALTA son las predominantes.

Con respecto al tipo de restricciones, es posible indicar que no existe una tendencia clara, siendo un tipo dominante, en cada variante encontrada (ver Tablas IX y X).

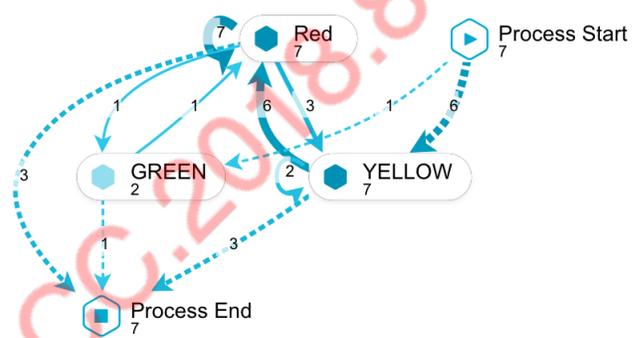


Figura 8. Proceso descubierto para los casos con loop RED-RED.

## V. TRABAJO RELACIONADO

La capacidad de procesar grandes volúmenes de información ha impulsado a diversas industrias a obtener nueva información valiosa respecto a los procesos involucrados. Este fenómeno, llamado *Big Data*, ha aparecido en diversas áreas siendo la construcción una de ellas. En la actualidad, se han desarrollado técnicas de *Data Mining* para aprender de proyectos pasados [18], hasta implementación de técnicas de *Machine Learning* y modelos de regresión para predecir tiempos de ciclo en actividades de construcción [19]. Sin embargo, una mirada desde la vista de *Process Mining* no ha sido estudiada [5]. *Process Mining* resulta ser la conexión directa entre *Data Science* (ej.: *Big Data*) y *Process Science*. Esta disciplina combina herramientas de las tecnologías de la información y de gestión con el fin de mejorar distintos procesos operacionales [6]. Su aplicación ha sido exitosa en diversas áreas, tales como procesos de negocios, salud [20], y asignación de recursos humanos [21]; áreas donde también se ha aplicado de forma exitosa la filosofía *Lean* [22]. Más detalles acerca de la relación entre los conceptos de *Process Science* y *Data Science* pueden ser encontrados en [9].

## VI. CONCLUSIONES Y TRABAJO FUTURO

Este trabajo ilustra el potencial de utilizar *Process Mining* en un contexto de la industria de la construcción, permitiendo concluir que es posible generar valor para este contexto, siempre y cuando, durante el proceso de *Process Mining*

Tabla IX  
CASOS POR TIPO DE RESTRICCIONES QUE TIENEN LOOPS RED-RED

Variante	%C.Previas	%M.O.	%Materiales	%Diseño	%Otros
Y-R-R-Y	83 %	0 %	17 %	0 %	0 %
Y-Y-R-R	0 %	0 %	100 %	0 %	0 %
G-R-R-Y	57 %	43 %	0 %	0 %	0 %
Y-R-R-R	0 %	43 %	43 %	0 %	14 %
Y-R-R-G	43 %	57 %	0 %	0 %	0 %

Tabla X  
CASOS POR IMPORTANCIA DE RESTRICCIONES QUE TIENEN LOOPS RED-RED

Variante	%ALTA	%MEDIAS	%BAJAS
Y-R-R-Y	83 %	17 %	0 %
G-R-R-Y	86 %	0 %	14 %
Y-R-R-R	100 %	0 %	0 %
Y-R-R-G	100 %	0 %	0 %
G-R-R-Y	86 %	0 %	14 %

se genere un diseño orientado a éste, que incluya todos los atributos necesarios para una sólida definición de casos y actividades. Es por eso que este artículo propone el *Constraint Bag Process Model* (CBPM), como una herramienta útil para tal análisis. El modelo generado, resulta ser una herramienta efectiva en la representación de casos particulares y generales desde una perspectiva de procesos, identificando cambios en el cumplimiento de restricciones semana a semana. La definición de actividades incluida en CBPM abre la posibilidad de ir descubriendo nuevos atributos relacionados en concordancia con lo buscado según el negocio y/o usuario; es decir, el modelo propuesto es altamente expansible y permite añadir diferentes atributos que permitan contestar diferentes preguntas. Queda como trabajo futuro, considerar nuevas perspectivas de análisis; por ejemplo, enriqueciendo la paleta de atributos descriptores propios del negocio y otros atributos calculados para estos propósitos. Además, es importante indicar que se puede hacer un estudio considerando los recursos (personas) así como su especialidad para analizar la incidencia de éstos en el comportamiento específico de las variantes detectadas.

#### AGRADECIMIENTOS

Los autores quieren agradecer a Celonis Academic por permitirnos utilizar la herramienta y al Centro de Excelencia de Gestión de Producción (GEPRO) por facilitarnos los datos disponibles en su software IMPERA. Este trabajo ha sido parcialmente financiado por el *Departamento de Ciencias de la Computación UC / Fond-DCC-2017-0001*.

#### REFERENCIAS

- [1] H. de Solminihaç and J. Dagá, "Productividad Media Laboral en la Construcción en Chile: Análisis Comparativo Internacional y con el Resto de la Economía," in *Centro Latinoamericano de Políticas Económicas y Sociales (CLAPES UC)*, Santiago, 2018, p. 23.
- [2] H. E. Picard, "Construction process measurement and improvement," in *10th Annual Conference of the International Group for Lean Construction*, C. T. Formoso and G. Ballard, Eds., Gramado, Brazil, 2002.
- [3] L. Koskela, G. Ballard, G. Howell, and I. Tommelein, "The foundations of lean construction," in *Design and Construction: Building in Value*, 2002.
- [4] G. Ballard and G. Howell, "Implementing Lean Construction: Stabilizing Work Flow," *Lean Construction*. A.A. Balkema, 1997. [Online]. Available: <http://www.leanconstruction.org>.
- [5] M. Bilal, L. O. Oyedele, J. Qadir, K. Munir, S. O. Ajayi, O. O. Akinade, H. A. Owolabi, H. A. Alaka, and M. Pasha, "Big Data in the construction industry: A review of present status, opportunities, and future trends," *Advanced Engineering Informatics*, vol. 30, no. 3, pp. 500–521, 2016.
- [6] W. M. P. van der Aalst, *Process Mining - Data Science in Action, Second Edition*. Springer, 2016.
- [7] W. M. van der Aalst, H. A. Reijers, A. J. Weijters, B. F. van Dongen, A. A. De Medeiros, M. Song, and H. Verbeek, "Business process mining: An industrial application," *Information Systems*, vol. 32, no. 5, pp. 713–732, 2007.
- [8] W. M. P. van der Aalst, A. Adriansyah, A. K. A. de Medeiros, F. Arcieri, T. Baier, T. Blickle, J. C. Bose, P. van den Brand, R. Brandtjen, J. Buijs, A. Burattin, J. Carmona, and et al., "Process mining manifesto," in *Business Process Management Workshops*. Springer Berlin Heidelberg, 2012, pp. 169–194.
- [9] W. van der Aalst and E. Damiani, "Processes meet big data: Connecting data science with process science," *IEEE Transactions on Services Computing*, vol. 8, no. 6, pp. 810–819, 2015.
- [10] G. Ballard and G. A. Howell, "An Update on Last Planner," in *11th Annual Conference of the International Group for Lean Construction*, 2003.
- [11] H. G. Ballard, "The Last Planner System of Production Control," Ph.D. dissertation, The University of Birmingham, 2000.
- [12] A. J. M. M. Weijters and J. T. S. Ribeiro, "Flexible heuristics miner (FHM)," in *Proceedings of the IEEE Symposium on Computational Intelligence and Data Mining, CIDM 2011, part of the IEEE Symposium Series on Computational Intelligence 2011, April 11-15, 2011, Paris, France*. IEEE, 2011, pp. 310–317.
- [13] S. J. J. Leemans, D. Fahland, and W. M. P. van der Aalst, "Process and deviation exploration with inductive visual miner," in *Proceedings of the BPM Demo Sessions 2014 Co-located with the 12th International Conference on Business Process Management (BPM 2014), Eindhoven, The Netherlands, September 10, 2014*, L. Limonad and B. Weber, Eds., vol. 1295, 2014, p. 46.
- [14] C. W. Günther and A. Rozinat, "Disco: Discover your processes," in *Proceedings of the Demonstration Track of the 10th International Conference on Business Process Management (BPM 2012), Tallinn, Estonia, September 4, 2012*, N. Lohmann and S. Moser, Eds., vol. 940, 2012, pp. 40–44.
- [15] F. Veit, J. Geyer-Klingenberg, J. Madrzak, M. Haug, and J. Thomson, "The proactive insights engine: Process mining meets machine learning and artificial intelligence," in *Proceedings of the BPM Demo Track and BPM Dissertation Award co-located with 15th International Conference on Business Process Modeling (BPM 2017)*,

- Barcelona, Spain, September 13, 2017., ser. CEUR Workshop Proceedings, R. Clarisó, H. Leopold, J. Mendling, W. M. P. van der Aalst, A. Kumar, B. T. Pentland, and M. Weske, Eds., vol. 1920. CEUR-WS.org, 2017. [Online]. Available: [http://ceur-ws.org/Vol-1920/BPM\\\_2017\\\_paper\\\_192.pdf](http://ceur-ws.org/Vol-1920/BPM\_2017\_paper\_192.pdf)
- [16] J. Geyer-Klingeberg, J. Nakladal, F. Baldauf, and F. Veit, "Process mining and robotic process automation: A perfect match;" in *Proceedings of the Dissertation Award, Demonstration, and Industrial Track at BPM 2018 co-located with 16th International Conference on Business Process Management (BPM 2018), Sydney, Australia, September 9-14, 2018.*, ser. CEUR Workshop Proceedings, W. M. P. van der Aalst, F. Casati, R. Conforti, M. de Leoni, M. Dumas, A. Kumar, J. Mendling, S. Nepal, B. T. Pentland, and B. Weber, Eds., vol. 2196. CEUR-WS.org, 2018, pp. 124–131. [Online]. Available: [http://ceur-ws.org/Vol-2196/BPM\\\_2018\\\_paper\\\_28.pdf](http://ceur-ws.org/Vol-2196/BPM\_2018\_paper\_28.pdf)
- [17] L. F. Alarcón and R. Calderón, "A production planning support system for construction projects," in *11th Annual Conference of the International Group for Lean Construction. Virginia, USA, 2003.*
- [18] P. Carrillo, J. Harding, and A. Choudhary, "Knowledge discovery from post-project reviews," *Construction Management and Economics*, vol. 29, no. 7, pp. 713–723, 2011.
- [19] M.-F. Siu, R. Ekyalimpa, M. Lu, and S. Abourizk, "Applying Regression Analysis to Predict and Classify Construction Cycle Time," in *Computing in Civil Engineering - Proceedings of the 2013 ASCE International Workshop on Computing in Civil Engineering*, 2013, pp. 669–676.
- [20] E. Rojas, J. Munoz-Gama, M. Sepúlveda, and D. Capurro, "Process mining in healthcare: A literature review," *Journal of Biomedical Informatics*, vol. 61, pp. 224–236, 2016.
- [21] M. Arias, R. Saavedra, M. R. Marques, J. Munoz-Gama, and M. Sepúlveda, "Human resource allocation in business process management and process mining: A systematic mapping study," *Management Decision*, vol. 56, no. 2, pp. 376–405, 2018.
- [22] A. D'Andreamatteo, L. Ianni, F. Lega, and M. Sargiacomo, "Lean in healthcare: A comprehensive review," *Health Policy*, vol. 119, no. 9, pp. 1197–1209, sep 2015.

preprint 10.1109/SCCC.2018.8705238