

TECNICHE DI PROGRAMMAZIONE CON VINCOLI PER LA VALIDAZIONE DI UNA BASE DATI SUL SISTEMA FOGNARIO.

Stefano BISTARELLI (*), Giovanni CIARDI(**), Franca MORONI(**)

(*) Dipartimento di Informatica, Università di Pisa , Corso Italia 40, Pisa – bista@di.unipi.it
(**) Servizio Sistemi Informativi Geografici, Regione Emilia-Romagna , v.le Silvani 4/3, Bologna
{ gciardi, fmoroni } @regione.emilia-romagna.it

Riassunto

Obiettivo del presente lavoro è l'illustrazione dell'applicazione di alcune tecniche di programmazione con vincoli per la verifica di consistenza delle informazioni contenute nella base dati: *"Base informativa delle reti fognarie della Regione Emilia-Romagna"*.

La base dati esistente registra: a) La configurazione spaziale dei tronchi e dei nodi tecnologici delle reti fognarie, b) Gli attributi fisici dei tronchi e dei nodi, c) Il verso del flusso nei tronchi della rete, d) Il funzionamento delle strutture tecnologiche specialistiche.

Le precedenti informazioni, utili ai fini della pianificazione e della programmazione delle reti tecnologiche, sono state ottenute mediante due campagne di rilevamento effettuate sulla base di un medesimo modello concettuale dei dati, ma realizzate da differenti soggetti istituzionali.

Se a ciò si aggiungono gli effetti della separazione fra fase di rilevamento e fase di caricamento dei dati, si comprende facilmente come un controllo di consistenza dell'informazione acquisita sia condizione preliminare per l'utilizzo della base dati medesima.

La formalizzazione delle relazioni tra attributi della rete in termini di vincolo ha consentito di individuare i domini assenti per le variabili che descrivono il sistema fisico, quindi di individuare a posteriori l'informazione non consistente. Costituisce un'estensione in corso dell'attività del gruppo di lavoro, la predisposizione di algoritmi che permettano solo il caricamento di informazioni coerenti con il sistema dei vincoli predeterminati.

Abstract

In this paper we show how Constraint Satisfaction Problems (CSPs) can be used to represent and efficiently detect inconsistency over a Data Base (DB). In this work we consider the DB representing the urban drainage systems of the Emilia Romagna region (Italy).

The DB contains geographic information of the arcs and of the nodes of the drainage network, The physic characteristics of arcs and nodes, the arc direction and a specification of the nodes w.r.t. their technologic use. The information present in the DB are obtained by different checking and in different moments. For this reason several data are misconfigured or inconsistent.

A first use of the constraint model we developed can be used to check consistency and to detect misconfigured data in the DB. The next step is to use the constraint definition as a filter to be imposed during the input of the data. Moreover the possibility to enhance the constraint with quality information (i.e. costs, temporal duration, etc.) could be useful to look for the "best" configuration of the network when more than one is possible.

Introduzione

Problemi di soddisfacimento di Vincoli

Un vincolo è una relazione logica fra una o più variabili, definite sui rispettivi domini. Il vincolo agisce restringendo il numero di valori del dominio che le variabili possono assumere. Con i termini: "problemi di soddisfacimento di Vincoli" (CSPs dall'inglese Constraint Satisfaction Problems) [Montanari U. (1974)] si fa riferimento a un formalismo generale con cui si possono

descrivere sistemi di oggetti reali e loro relazioni. Nel nostro caso tali relazioni rappresentano specifiche proprietà degli oggetti della base di dati. CSPs per cui il dominio delle variabili è finito assumono uno speciale interesse applicativo perché più facili da risolvere e al tempo stesso ancora sufficientemente espressivi per rappresentare situazioni di vita reale.

Un CSP su dominio finito è una tupla $\langle V, D, C, con, def \rangle$ dove

- V è un insieme finito di variabili $V = \{v_1, \dots, v_n\}$;
- D è un insieme finito di valori detto *Dominio*;
- C è un insieme finito di vincoli $C = \{c_1, \dots, c_m\}$;
- *con* è detta funzione di *connessione* e specifica le variabili coinvolte in ogni vincolo;
- *def* è detta funzione di *definizione* e specifica i valori ammessi per le variabili di ciascun vincolo.

Un CSP può facilmente essere rappresentato come un ipergrafo etichettato rappresentando le variabili come nodi e i vincoli come archi che connettono le variabili coinvolte nel vincolo (date dalla funzione *con*).

Ogni arco sarà etichettato con le tuple di valori del dominio permessi per le variabili coinvolte (e specificate dalla funzione *def*). Come esempio consideriamo il CSP rappresentato in Figura 1:

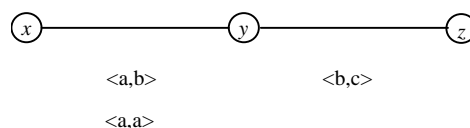


Figura 1: un esempio di CPS.

Nell'esempio si hanno 3 variabili $V = \{x, y, z\}$ i cui possibili valori sono $D = \{a, b, c\}$.

Sono presenti due vincoli: uno tra le variabili x e y che permette le tuple $\langle a, b \rangle$ e $\langle a, a \rangle$, e un secondo tra le variabili y e z che permette la tupla $\langle b, c \rangle$.

Risolvere un problema di vincoli significa trovare una istanziazione per tutte le variabili che soddisfi tutti i vincoli. È facile verificare che nel CSP in Figura 1 l'unica soluzione possibile è rappresentata dalla tupla $\langle a, b, c \rangle$ che assegna rispettivamente $x = a$, $y = b$ e $z = c$.

La base informativa delle reti fognarie

Nel presente lavoro si è fatto ricorso alla programmazione con vincoli per ripristinare la coerenza interna dei dati implementati nella "*Base informativa delle reti fognarie*" della Regione Emilia-Romagna. Tale base di dati è stata progettata nel 1992. Negli anni seguenti, mediante due separate campagne di rilevamento, si sono prodotti: a) le carte tematiche contenenti la geometria della rete (su fondo topografico C.T.R. 1:10.000), b) le schede contenenti le informazioni sugli aspetti costruttivi delle condotte e sui dettagli tecnologici degli impianti localizzati (depuratori, scaricatori di piena, impianti di sollevamento ecc.).

Le fasi di progettazione, rilevamento, realizzazione della cartografia tematica, caricamento dei dati e controllo sono state gestite da soggetti istituzionali differenti.

Di qui la necessità di svolgere un controllo sulla coerenza semantica delle informazioni e di ripristinare a posteriori, quella consistenza dei dati supporto necessario alla validità di modelli di previsione del comportamento fisico delle reti fognarie.

Vincoli sulla base dati

Vincolo Intra-nodo

Le reti fognarie sono schematizzate mediante una topologia arco-nodo.

I nodi della rete sono di due tipi: a) nodi NO: rappresentano il luogo ove si verifica una discontinuità nei parametri che caratterizzano un tronco (p. es. variazione di sezione, di materiale, ecc.), b) nodi che rappresentano il luogo ove si collocano uno o più impianti tecnologici (p. es. depuratori (ID), impianti di sollevamento (IS), ecc.)

Il corretto funzionamento della rete esclude la possibilità che alcuni impianti coesistano nello stesso luogo. Nella base di dati la suddetta condizione è esprimibile mediante un vincolo che agisce sul campo "tipo", campo che contiene i descrittori degli apparati tecnici presenti nel nodo. Questo tipo di vincolo è detto unario in quanto agisce sul dominio di un'unica variabile.

L'insieme delle funzioni tecniche elementari svolte dai nodi della rete è :

$T=\{0, NO, ID, IF, IS, SP, SI, IP, DS, IR, RI, SR, RS, VA\}$ dove: 0 - non conosciuto; NO - nodo ordinario; ID - Impianti di depurazione; IF - immissione finale di scarico di depuratore in acque superficiali; IS - impianto di sollevamento; SP - scaricatore di piena; SI - immissione finale di scaricatore di piena in acque superficiali; IP - immissione finale di scarico non depurato in acque superficiali; DS - dispersione di scarico non depurato in terreno superficiale; IR - immissione di scarico non depurato in altra rete fognaria; RI - ricevimento di scarico non depurato da altra rete fognaria; SR - immissione di acque sfiorate in altra rete fognaria; RS - ricevimento di acque sfiorate da altra rete fognaria; VA - vasca di accumulo e/o di pioggia.

Si costruisce un vincolo unario sulla variabile *tipo* con dominio $D = \wp(T)$, (dove $\wp(T)$ è l'insieme dei sottoinsiemi di T). La Tabella seguente definisce le funzioni tecniche del vincolo intra-nodo rappresentate dalla funzione *def.* introdotta nel paragrafo: problemi di soddisfacimento di vincoli

<i>tipo</i>	0	NO	ID	IF	IS	SP	SI	IP	DS	IR	RI	SR	RS	VA
0		no	no	no	no	no	no	no	no	no	no	no	no	no
NO	no		no	no	no	no	no	no	no	no	no	no	no	no
ID	no	no		si	si	si	no	no	no	no	no	no	no	no
IF	no	no	si		si	no	no	no	no	no	no	no	no	no
IS	no	no	si	si		si	no	no	no	no	no	no	no	no
SP	no	no	si	si	si		no	no	no	no	no	no	no	no
SI	no	no	no	no	si	no		no	no	no	no	no	no	si
IP	no	no	no	no	si	no	no		no	no	no	no	no	no
DS	no	no	no	no	no	no	no	no		no	no	no	si	si
IR	no	no	no	no	si	no	no	no	no		no	no	no	no
RI	no	no	no	no	no	no	no	no	no	no		no	si	si
SR	no	no	no	no	no	no	no	no	no	no	no		si	no
RS	no	no	no	no	no	no	no	no	no	no	no	si		no
VA	no	no	no	no	si	si	no	no	no	si	no	no	no	

Tabella 2: funzione di definizione: *funzioni_tecniche*.

Graficamente tale vincolo unario sarà rappresentato come in Figura 3.

Vincoli Intra-tronco

Ad ogni tronco della rete sono associati attributi che descrivono le caratteristiche del flusso e attributi che descrivono le caratteristiche costruttive di esso. La relazione fra il tipo di flusso e la pressione in un tronco è descritta da un vincolo di tipo binario che agisce su due campi della base dati. Le variabili ed i rispettivi domini sono: $V=\{tip_flusso; press\}$ e $D =\{pelo\ libero, pressione, cielo\ aperto\}$ e $D=\{0, GT.0^1\}$ rispettivamente.

La funzione di definizione *caratteristiche_flusso*, è data dalle coppie ammissibili:

$$\{pelo\ libero, 0\}, \{pressione, GT, 0\}, \{cielo\ aperto, 0\},$$

¹ GT.0 indica qualsiasi valore della pressione $\neq 0$.

In Figura 4 è rappresentato il vincolo binario flusso nel tronco.

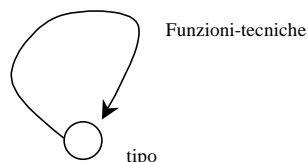


Figura 3: CSP nodi.

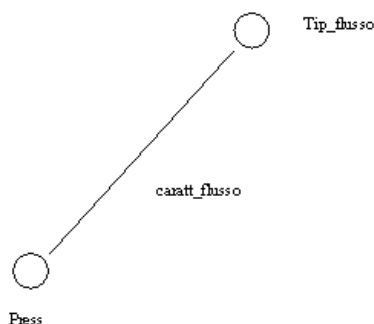


Figura 4: CPS flusso nel tronco.

Il secondo sistema di vincoli esprime, invece, le relazioni fra le caratteristiche strutturali dei tronchi. Le variabili e i loro domini sono, rispettivamente:

$V = \{Sez, Diam, Larg, Mat, Riv\}$ e

$D_{Sez} = \{\text{circolare, ovoidale, vigentina, vigentina con cunetta, scatolare, scatolare con cunetta, semi ovoidale, canaletta, trapezoidale chiusa, trapezoidale aperta, canaletta con cunetta}\}$,

$D_{Diam} = \{250, 3800\}$,

$D_{Larg} = \{250, 3105\}$,

$D_{Mat} = \{\text{cls. pref. vibrato, cls. pref. centrifugato, cls pref. precompresso, cls.in opera, murata laterizio, cemento amianto, gres ceramico, metallo, PVC}\}$,

$D_{Riv} = \{\text{grezzo intonacato, fondello gres, piastrelle gres, resine epossidiche o epossicatramose, silicatizzanti, vetrificanti, bituminosi}\}$.

dove: *Sez* = forma della sezione del tronco; *Diam* = rispettivamente dimensione verticale o diametro della sezione (sezioni circolari); *Larg* = dimensione orizzontale della sezione; sono le variabili che esprimono la geometria del tronco; *Mat* = materiale della condotta; *Riv* = tipo di rivestimento.

Le due funzioni di definizione sono rispettivamente: *caratter_costruttive1* (il rivestimento non pone alcun tipo di vincolo).

<i>Mat</i>	<i>Sez</i>	Circolare	Ovoidale	Vicentina	Vicentina cun.ta	Scatolare	Scatolare cun.ta	Semi ovoidale	Canaletta	Trapezia chiusa	Trapezia aperta	Canaletta Cun.ta
Cls. vibrato		si	si	si	si	si	si	si	si	si	si	si
Cls. centrif.		si	si	si	si	si	si	si	si	si	si	si
Cls. prec.so		si	si	si	si	si	si	si	si	si	si	si
Cls.in opera		no	no	no	no	no	no	no	no	no	no	no
Muratura lat.		no	no	si	si	si	si	no	si	si	si	no
Cem.to_am.to		si	si	si	si	si	si	si	si	si	si	si
Gres-cer.co		si	no	no	no	no	no	no	no	no	no	no
Metallo		si	no	no	no	no	no	no	no	no	no	no
P.V.C.		si	no	no	no	no	no	no	no	no	no	no

Tabella 5 - Funzione descrittiva *caratteristiche-costruttive1*.

e *caratteristiche_costruttive2* (il materiale di rivestimento concorre a definire il vincolo).

Costituiscono valori assenti della terna: $\{Sez, Mat, Rivest\}$:

<i>Sez</i>	<i>Mat</i>	<i>Rivest</i>
circolare	gres	Grezzo
circolare	metallo	Grezzo
circolare	pvc	Grezzo
circolare	metallo	Bituminoso
circolare	pvc	Bituminoso

Tabella 6: Funzione descrittiva *caratt_costruttive2*.

Il vincolo *caratt_costruttive2* è un vincolo di tipo ternario (coinvolge cioè 3 variabili). Nella figura che segue è rappresentata il CSP per i due vincoli appena descritti.

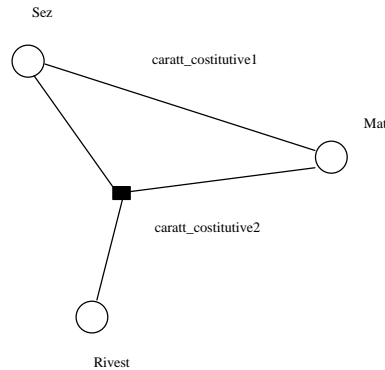


Figura 7: CSP vincoli intratronco.

Vincoli Tronco-tronco

Perché la circolazione del flusso sia garantita, è necessario che non siano violate le condizioni che consentono la continuità del flusso e rappresentate dalla equazione $\sum_{i=1}^N Q_i^+ \leq \sum_{i=1}^M Q_i^-$, dove: Q_i sono le portate massime nominali nella sezione di tronco coincidente con il nodo rispetto a cui si scrive la continuità; il segno + o - indica rispettivamente il verso del flusso nel nodo. Nella modellizzazione tutto ciò è stato semplificato sostituendo alle sommatorie delle portate in entrata e in uscita rispetto al nodo di riferimento, le portate nominali di due soli tronchi fittizi ipotizzati a monte e a valle del nodo. Possiamo quindi considerare le seguenti variabili: $\mathbf{V}=\{Q_S, Q_D, tip_flusso_S, tip_flusso_D\}$ e i corrispondenti domini: $\mathbf{D}_Q=[Q_{min}, Q_{max}]$ e $\mathbf{D}_{tip-flusso}=\{\text{pelo libero, pressione, cielo aperto}\}$. Se la variabile *tipo_flusso* assume il valore *cielo aperto* o *pelo libero*, la portata del tronco sorgente dovrà essere strettamente minore di quella del tronco di destinazione.

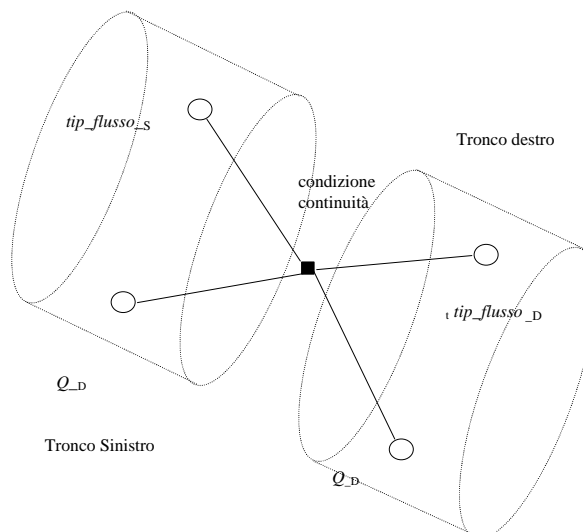


Figura 8: CSP condizione di continuità

Il CSP risultante è rappresentato in Figura 8.

Vincolo Tronco-nodo

Questo tipo di vincolo garantisce la corretta giustapposizione degli apparati tecnologici della rete. Ad ogni tipo di nodo deve seguire un tronco caratterizzato da un flusso e una tipologia del fluido veicolato ad esso coerenti.

Le variabili ed i loro domini sono $V = \{tipo, tipo_uso, tipo_flusso\}$ dove D_{tipo} e D_{tipo_flusso} sono già quelli precedentemente definiti e $D_{tipo_uso} = \{bianche, nere miste, scarico di depuratore in corpo idrico, da scaricatore di piena\}$.

In questo caso le triple che definiscono il vincolo, rispetto al nodo ID sono:

$def = \{ \langle ID, \text{recapito da depuratore in corpo idrico, pelo libero} \rangle, \langle ID, \text{recapito da depuratore in corpo idrico, pressione} \rangle, \langle ID, \text{recapito da depuratore in corpo idrico, cielo aperto} \rangle, \langle ID, \text{recapito da depuratore in corpo idrico, pelo libero} \rangle, \langle ID, \text{recapito da depuratore in corpo idrico, pelo libero} \rangle \}$.

La Figura seguente mostra il vincolo ternario sopra descritto.

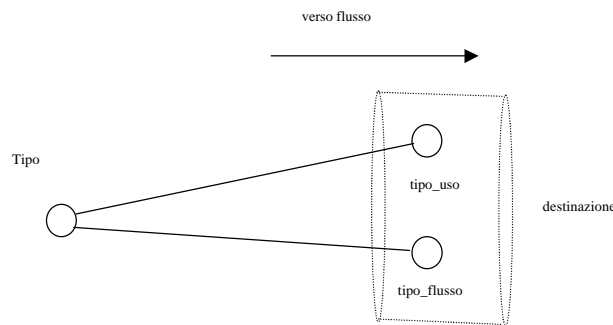


Figura 9: CSP nodo-tronco

Conclusioni e sviluppi

Una possibile estensione del lavoro fatto finora sarà l'utilizzazione di un *soft CSP* [Bistarelli S. (2001)] per la rappresentazione dei vincoli della base di dati. In questo modo si potranno arricchire i vincoli con etichette qualitative e quantitative al fine di catturare maggiore informazioni dal data base.

Possibili esempi sono l'assegnazione di costi ai vari elementi strutturali dei componenti la rete fognaria (tronchi e nodi) al fine di avere a disposizione non solo una informazione sulle possibile configurazioni corrette, ma su un loro costo.

Similmente, assegnando etichette indicanti la durata temporale del materiale, si potranno avere informazioni sulla durata complessiva di ogni singola sottorete (intervalli di manutenzione). Tale approccio è in studio con partner internazionali.

Bibliografia

- Bistarelli S. (2001) – *Soft Constraint Solving and Programming : A general framework*. Phd thesis, Dipartimento di Informatica, Università di pisa, 2001
- Montanari U. (1974) – *Networks of constraints: Fundamental properties and applications to picture processing*. Information Science, 7:95-132, 1974.