Seventh Int. Workshop on GNs, Sofia, 14-15 July 2006, 15-19

GENERALIZED NET MODEL OF MULTI-SOURCE DATABASE SYSTEM WITH DIFFERENT ACCESS TIMES

Boyan Kolev [1], Evdokiya Sotirova[2], Krassimir Atanassov[1] and Panagiotis Chountas[3]

[1] CLBME - Bulgarian Academy of Sciences, Bl. 105, Sofia- 1113, Bulgaria, e-mails: *bik@clbme.bas.bg* and *krat@bas.bg*

[2] Prof. Asen Zlatarov University, Bourgas 8000, Bulgaria, e-mail: *esotirova@btu.bg*

[4] University of Westminster-Health Care Computing Group, HSCS, Northwick Park, London, HA1 3TP, UK, e-mail: *chountp@wmin.ac.uk*

1 Introduction

In this paper we are concerned with the query of an integrated multi- source environment, with the emphasis on determining the source S that can satisfy a query Q *while* the response time Δt for retrieving the answer is kept to minimal. To this extent if there are more one sources that could possibly satisfy Q we are interesting in the source that can provide a satisfactory answer at a time $t \leq \Delta t$.

Given the response time Δt we can establish a partial order for the underlying sources $S_1, ..., S_n$ with respect to query Q. Thus the underlying source S_i with relevant facts $f_1, ..., f_n$ that satisfy a query Q and achieves response time $t \leq \Delta t$ is given the highest priority always with respect to query Q for $1 \leq i \leq n$.

In defining a common integrated info-base, we assume that there exists a single (hypothetical) info-base that represents the real world. This ideal info-base includes only perfect descriptions.

We now formulate two assumptions for the governing purposes of the hypothetical info-base [5]. These assumptions are similar to the Universal Scheme Assumption and the Universal. These two assumptions are statements of reconciliation of the underlying Info-bases.

The Scheme Soundness Principle (SSP). All conceptual schemes are derivatives of the real world scheme. That is, in each conceptual scheme, every structural component is a view of the real world scheme. The meaning of this assumption is that the different ways in which reality is modelled are all correct; i.e., there are no modelling errors, only modelling differences. To put it in yet a different way, all intentional inconsistencies among the independent conceptual schemes are reconcilable.

The Instance Soundness Principle (ISP). All database-fact instances are derivatives of the real world instance. The meaning of this assumption is that the information stored in info-bases is always correct; i.e., there is no erroneous information, only different semantic representations of alike facts.

It is suggested by that in query-answering systems one central task is to compare two information items, one from the client and the other from the info-base. In our framework a client can be defined as an application that is trying to build its own info-base out of the existing reconcileable info-bases, or an unaware hypothetical user trying to extract information similar to its request from the info-base, that is the closest one to the single (hypothetical-perfect) database, which represents the real world according to **SSP, ISP** principles.

2 GN-model of multi-source database system

Let us have s different DataBases (DBs), that have different time periods of work. We shall interprete these DBs by tokens $\delta_1, ..., \delta_s$.

Let there be an Expert System (ES) and a System Administrator (AS) that we shall interprete by token β. Let a Special Module (SM), interpreted by token γ, control the process of changing of information from one DB to another one. Let requests (we shall mark all them by tokens α) enter ES and let them have user's preference for the duration of the time for obtaining of the answer. Therefore, the initial characteristic of α-token is

"request; user's preference for the duration of the time".

Initially, tokens $\delta_1, ..., \delta_s$ stay in place l_9 with initial characteristics

"i-th DB; time for access",

where $1 \leq i \leq s$. Token β permanently stays in place l_4 with initial characteristic

"ES and AS; specific parameters".

Finally, token γ permanently stays in place l_{10} with initial characteristic

"SM; specific parameters; C_1, C_2",

where C_1 and C_2 are constants.

The GN contains 5 transitions with the following forms.

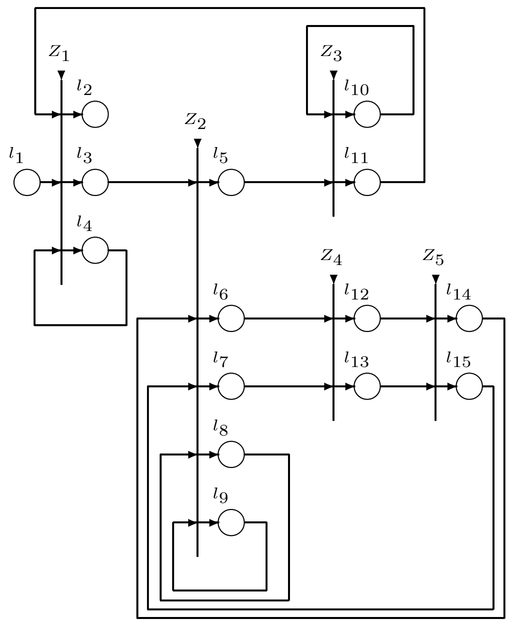
$$Z_1 =< \{l_1, l_4, l_{11}\}, \{l_2, l_3, l_4\}, \begin{array}{c|ccc} & l_2 & l_3 & l_4 \\ \hline l_1 & false & false & true \\ l_4 & false & W_{4,3} & true \\ l_{11} & true & false & false \end{array} >,$$

where
$W_{4,3} =$ "there is a DB that contains the necessary information".

The α-token enters place l_2 without a new characteristic, while in place l_3 it obtains the characteristic

"the necessary information is contained in k-th DB, where $1 \leq k \leq s$".

The β-token does not obtain any new characteristic.



$$Z_2 =< \{l_3, l_8, l_9, l_{14}, l_{15}\}, \{l_5, l_6, l_7, l_8, l_9\}, \begin{array}{c|ccccc} & l_5 & l_6 & l_7 & l_8 & l_9 \\ \hline l_3 & true & false & false & false & false \\ l_8 & false & false & false & false & true \\ l_9 & false & W_{9,6} & W_{9,7} & W_{9,8} & true \\ l_{14} & false & false & false & false & true \\ l_{15} & false & false & false & false & true \end{array} >,$$

where
$W_{9,6} =$ "token γ;s last characteristic is the number of a the current DB",
$W_{9,7} =$ "the current DB's access is higher that the average access time",

$W_{9,8} =$ "the current DB's contains the necessary information for token α".
In place l_5 the α-token obtains the characteristics

"answer of the question".

Each of the δ-tokens can split to two or three tokens - the same δ-token that continues to stay in place l_9, and either to token δ_i^* that enters place l_8 without a new characteristic, where i is the number of the respective δ-token; or to tokens δ_i^1 and δ_i^2 that enter places l_6 and l_7, respectively, where i is used in the above sense. Token δ_i^1 obtains the characteristic

"localisation of the information that must be deleted from the DB"

in place l_6, and the tokens do not obtain any characteristics in places l_7 and l_9.

$$Z_3 =< \{l_5, l_{10}\}, \{l_{10}, l_{11}\}, \begin{array}{c|cc} & l_{10} & l_{11} \\ \hline l_5 & false & true \\ l_{10} & true & false \end{array} > .$$

In place l_{11} the α-token obtains the characteristics

"duration of obtaining of the transfer",

while in place l_{10} the γ-token, which should have smaller priority than token α, obtains the characteristic

"number of cases, when the access to the present information is longer than the access, requested by the user".

$$Z_4 =< \{l_6, l_7\}, \{l_{12}, l_{13}\}, \begin{array}{c|cc} & l_{12} & l_{13} \\ \hline l_6 & true & false \\ l_7 & false & true \end{array} > .$$

The δ^1-token obtains the characteristic

"cleaning of the information from the current DB"

in place l_{12}, while the δ^2-token does not obtain any characteristic in place l_{13}.

$$Z_5 =< \{l_{12}, l_{13}\}, \{l_{14}, l_{15}\}, \begin{array}{c|cc} & l_{14} & l_{15} \\ \hline l_{12} & true & false \\ l_{13} & false & true \end{array} > .$$

The δ^2-token obtains the characteristics

"record of the information from the DB represented by token δ^1"

in place l_{15}, while the δ^1-token does not obtain any characteristic in place l_{14}.

3 Conclusion

In this paper we analysed the issue of establishing a partial order for the underlying sources $S_1, ..., S_n$ with respect to query Q. Thus the underlying source S_i that satisfy a query Q and achieves response time $t \leq \Delta t$ is given the highest priority always with respect to query Q for $1 \leq i \leq n$.

In enhancing our initial approach we need to analyse the issue of *"approximate answers"*. To put it differently, a query Q may have several candidate answers. In such cases the most authoritative answer must be retrieved with respect to the perfect answer.

If new sources of information need to be incorporated in the integrated environment then the partial order for the underlying sources $S_1, ..., S_n$ with respect to query Q needs to be recalculated.

In future, we are outlying some thoughts that will enable us to represent conflicting information as part of a 4 valued characteristic function of para-consistent relation, which maps tuples to one of the following values: $\top$ (for contradiction), t (for true), f (for false) and $\perp$ (for unknown). This will let to reason about conflict a quantitative four- value logic. The elements of the temporal information (temporal facts) can be represented in the form $\langle F, t_L^F, t_R^F \rangle$, where $[t_L^F, t_R^F]$ is a time interval. Using the ideas for Intuitionistic fuzzy expert systems [3, 4], we can estimate any fact F and it can obtain Intuitionistic fuzzy truth-values $V(F) = \langle \mu^F, \nu^F \rangle$, such that $\mu^F, \nu^F \in [0, 1]$ and $\mu^F + \nu^F \leq 1$. Therefore, the above fact can be represented in the form

$$\langle F, t_L^F, t_R^F, \mu^F, \nu^F \rangle.$$

This form of the fact corresponds to the case in which the fact is valid in interval $t^F = [t_L^F, t_R^F]$ and at every moment of that interval the fact has the truth-value $\langle \mu^F, \nu^F \rangle$. Thus it is required to develop a relational Intuitionistic environment for representing uncertainty and contradiction or conflict as part of a multi-source database environment.

The paper is a continuation of the ideas from [2, 6].

References

[1] Atanassov K. *Generalized Nets.* Singapore, New Jersey, London, World Scientific, 1991.

[2] Atanassov, K. Generalized Nets in Artificial Intelligence. Vol. 1: Generalized nets and Expert Systems. "Prof. M. Drinov" Academic Publishing House, Sofia, 1998.

[3] Atanassov, K. Intuitionistic Fuzzy Sets. Springer Physica-Verlag, Heidelberg, 1999.

[4] Atanassov K., Temporal Intuitionistic fuzzy relations. Proceedings of the Fourth International Conference on Flexible Query Answering Systems, FQAS 2000, Advances in Soft Computing, ISBN 3-7908- 1347-8, Physica-Verlag, 153-160, 2000.

[5] Chountas, P, Petrounias, I., Representation of Definite, Indefinite and Infinite Temporal Information, Proceedings of the 4th International Database Engineering & Applications Symposium (IDEAS'2000), IEEE Computer Society, ISBN 0-7695-0789-1, 167-178, 2000.

[6] Kolev, B., E. El-Darzi, E. Sotirova, I. Petronias, K. Atanassov, P. Chountas, V. Kodogianis. Generalized Nets in Artificial Intelligence. Vol. 3: Generalized nets, Relational Data Bases and Expert Systems. "Prof. M. Drinov" Academic Publishing House, Sofia, 2006.