

# Creating Quantitative Goal Models: Governmental Experience

Okhaide Akhigbe<sup>1</sup>, Mohammad Alhaj<sup>1</sup>, Daniel Amyot<sup>1</sup>, Omar Badreddin<sup>2</sup>,  
Edna Braun<sup>1</sup>, Nick Cartwright<sup>1</sup>, Gregory Richards<sup>1</sup>, and Gunter Mussbacher<sup>3</sup>

<sup>1</sup>University of Ottawa, <sup>2</sup>Northern Arizona University, <sup>3</sup>McGill University  
okhaide@uottawa.ca, malhaj@sce.carleton.ca,  
damyot@eecs.uottawa.ca, Omar.Badreddin@nau.edu,  
eawbraun@gmail.com, ncart@sympatico.ca,  
richards@telfer.uottawa.ca, gunter.mussbacher@mcgill.ca

**Abstract.** Precision in goal models can be enhanced using quantitative rather than qualitative scales. Selecting appropriate values is however often difficult, especially when groups of stakeholders are involved. This paper identifies and compares generic and domain-specific group decision approaches for selecting quantitative values in goal models. It then reports on the use of two approaches targeting quantitative contributions, actor importance, and indicator definitions in the Goal-oriented Requirement Language. The approaches have been deployed in two independent branches of the Canadian government.

**Keywords:** AHP, Compliance, Contributions, Decision Making, Enterprise Architecture, GRL, Indicators, Quantitative Values.

## 1 Introduction

Goal modeling offers a way of structuring requirements according to their contribution towards achieving the objectives of various stakeholders. Goals can be decomposed and linked, and trade-offs can be evaluated when stakeholder objectives are conflicting. Common goal modeling languages include *i\** [28], Tropos [7], KAOS [25], the Goal-oriented Requirement Language (GRL) [11], and the Business Intelligence Model (BIM) [10]. Each language comes with different sets of concepts and analytic capabilities. Most languages however have a *contribution* concept, used to indicate how much a goal model element influences, positively or negatively, another model element. *Qualitative* contribution scales often specify the level of sufficiency of the positive/negative contribution (sufficient, insufficient, or unknown), leading to a handful of possible contribution value combinations. Such coarse-grained qualitative scales are useful in a context where little information is known about the domain, or when there is uncertainty about the exact degree of contribution.

A recent trend in such languages is to support *quantitative* contribution scales, with numerical values. Such finer-grained scales enable modelers and analysts to better differentiate between the contributions of alternatives to higher-level objectives. However, whereas agreeing on the positive negative nature of a contribution is often

easy, deciding on valid contribution/importance values on a quantitative scale is not trivial, especially when different groups of stakeholders are involved.

There are many generic and domain-specific group decision-making approaches. This paper presents our experience using some of these approaches to create quantitative GRL goal models at two different departments of the Canadian government. Section 2 provides background on GRL and defines some key quantification challenges. Section 3 identifies seven generic and domain-specific group decision approaches for selecting quantitative values in goal models. Section 4 presents our experience using consensus to derive indicators from regulations and generating legal models for compliance, while Section 5 illustrates our experience using the Analytic Hierarchy Process (AHP) for contribution, importance and indicator values in enterprise architecture goal models. Section 6 discusses lessons learned and Section 7 our conclusions.

## 2 Background and Goal Modeling Challenges

Both quantitative and qualitative techniques in goal modeling are accommodated adequately in the Goal-oriented Requirement Language – GRL, one of the sub-languages of the User Requirements Notation (URN), whose standard was revised in 2012 [11]. GRL has four primary elements: *intentional elements* (goals, softgoals, tasks, resources, and beliefs), *indicators*, *intentional links*, and *actors* (basically various forms of stakeholders, or the system itself, which contain intentional elements).

*Goals* can be achieved fully whereas *softgoals* are usually satisfied (or “satisfied”) to a suitable extent. *Tasks* represent solutions and may require *resources* to be utilized [4]. Intentional links connect elements in a goal model: *decompositions* allow elements to be decomposed into sub-elements through AND/OR/XOR relationships, *contributions* model desired impacts of elements on other elements qualitatively or quantitatively, *correlations* describe side effects rather than impacts, and *dependencies* model relationships between actors. The quantitative scale used in GRL to describe contribution weights goes from –100 (break) to +100 (make). For simplicity in this paper, whatever applies to contributions also applies to correlations.

In GRL, a *strategy* captures a particular configuration of alternatives and satisfaction values in the GRL model by assigning an initial qualitative or quantitative satisfaction level to some intentional elements in the model. A GRL *evaluation algorithm* disseminates this information to other intentional elements of the model through their links and computes their satisfaction level [3]. Strategies can be compared to each other to facilitate the identification of the most appropriate trade-off amongst conflicting stakeholders’ goals. When evaluating the degree of satisfaction of goals, the qualitative approach uses measures such as “satisfied”, “partially satisfied”, “denied”, or “undetermined”. This scale may be too coarse-grained or vague in some decision-making contexts, and hence GRL also supports a *quantitative scale* that ranges from –100 (denied) to 0 (neutral) to +100 (satisfied). jUCMNav [5], an Eclipse tool for URN modeling and analysis, further supports a satisfaction scale more intuitive in some application areas, ranging from 0 (denied) to 50 (neutral) to 100 (satisfied).

Indicators allow real-life values to be the basis for goal model evaluations. A conversion function translates the real-world value of the indicator into a GRL quantitative satisfaction value. This conversion can be done through linear interpolation by comparing the real-life value against target, threshold, and worst-case values for that indicator, or through an explicit conversion table. Real-life values in indicators can be fed manually in a strategy (to explore what-if scenarios) or from external data sources (e.g., sensors or BI systems), turning the GRL model into a monitoring system.

GRL was selected in the two government projects presented here because 1) it is a standardized modeling language (a genuine concern for government agencies), 2) it supports quantitative evaluations, 3) it supports strategies and numerous propagation algorithms, with color feedback, 4) it supports indicators, which can be fed from external data sources, 5) it can be profiled to a domain (through URN metadata and URN links) while remaining within the boundaries of the standard, 6) good tool support is available (jUCMNav), 7) GRL had been successfully used in many projects, and 8) local expertise was readily available at the time these projects were done.

Like many other goal-oriented modeling languages, GRL has several limitations with respect to the use of decomposition links [21, 22], the lack of modularity [19, 22], or the cognitive fitness of its graphical syntax [20]. However, the main challenges of interest in this paper relate to the quantification of the model (contributions, indicators, and importance values): “what do the numbers mean” or “where are the numbers coming from” [14, 22]. Very often in GRL, it is more important to compare evaluation results of different strategies when making a decision than to focus on the exactitude of the values of a single evaluation. Also, decisions are not always very sensitive to values, and this is why jUCMNav also supports value ranges ([min, max]) for sensitivity analysis [5]. Yet, precision in quantitative values is still desirable, especially when multiple stakeholders are involved.

For the creation of goal models, there is an emerging trend in trying to use systematic approaches for reaching agreements on quantitative values, especially for goal contributions. In particular, the Analytic Hierarchy Process (AHP) technique [23] is used to take into consideration the opinions of many stakeholders, through surveys based on pairwise comparisons. In the last few years, relative contribution weights have hence been computed with AHP for  $i^*$  models by Liaskos et al. [16], for models using the Non-Functional Requirements framework by Kassab [13], and for models in a proprietary goal modeling language by Vinay et al. [27]. In all cases, these approaches targeted relative contributions, without support for negative contributions, for under-contributions (sum of relative contributions to a target element being less than 100%), or for over-contributions (sum greater than 100%). Although an AHP-based approach is useful when constructing goal models, it does not necessarily eliminate the need for validation and conflict detection, e.g., with questionnaires [9, 12].

When compared to related work, this paper focuses on the creation of quantitative goal models in a different and standardized language (GRL). Not only are contributions covered, but so are indicators and the importance of intentional elements. This paper also reports on the industrial application of two group decision approaches in two different departments of the Government of Canada, in real projects.

### 3 Group Decision Approaches for Goal Modeling

We have identified and evaluated seven approaches divided into two groups: generic and domain specific. Other approaches exist to support group decisions or multiple criteria decision analysis [26]. However, after a first filtering exercise, we limited ourselves to those that are suitable in a GRL context (e.g., without fuzzy values). The first five approaches are generic whereas the last two are specific to GRL models in the domain of regulatory compliance, which was studied in one of our two projects:

- **Equal Relative Weights (ERW):** All contributions targeting the same intentional element have equal weights, neglecting the fact that some contributors might be more important than others. This approach does not require any discussion.
- **Round-Table Discussion and Consensus (RTD&C):** A focus group method where experts are assembled in a dialog setting. Groupings of related choices, contained in models, are put up on a screen, and the experts are asked to discuss and assign relative weights to each choice in each grouping.
- **Delphi Process (DP)** [17, 18]: A method used to reach consensus amongst a group of experts. Participants answer short questionnaires and provide reasons for their answers through several rounds. After each round, an anonymous summary of the responses is provided to the participants.
- **Analytic Hierarchy Process (AHP)** [23]: A structured technique for organizing and analyzing complex decisions using pair-wise comparisons.
- **Approximate Weighting Method – Rank Order Centroid (ROC) Weights** [2]: Objectives are ordered from most to least important, and a number of different formulas are used to assign relative weights.
- **Relative Weights Derived from Regulatory Penalties (RP):** Relative weights for regulations within the models are assigned according to the *penalties* attached to each regulation (i.e., the more severe the penalty associated with a violation of a regulation, the higher it will score in terms of contribution).
- **Relative Weights Derived from Frequency of Inspection Requirements (FIR):** Relative weights for regulations within the models are assigned according to the *frequency* at which inspectors are required to check for compliance with a given regulation, and to the level of investigation associated with each regulation.

These seven approaches were evaluated according to criteria relevant in our context (Table 1). RTD&C and DP require face-to-face meetings (FF), while the other approaches do not (AHP can be done in a face-to-face or virtual meeting (VM), or through surveys). RTD&C, DP, and AHP also allow group thinking and require preparation. RTD&C is more susceptible to peer pressure. All approaches allow for record keeping. The last three criteria were all assessed qualitatively using High, Medium, or Low. In general, the assessment divides the seven approaches into two groups with similar results. The first group (RTD&C, DP, and AHP) requires high or medium preparation and meeting time, yields high accuracy, but offers low precision. The second group needs low preparation and meeting time, but is not as accurate even though precision is high. The ERW approach has the lowest accuracy.

**Table 1.** Comparison of group decision approaches

| Criteria       | Group Decision Approaches |      |          |      |        |        |        |
|----------------|---------------------------|------|----------|------|--------|--------|--------|
|                | RTD&C                     | DP   | AHP      | ERW  | ROC    | RP     | FIR    |
| Setting        | FF                        | FF   | FF/VM    | VM   | VM     | VM     | VM     |
| Peer Pressure  | Yes                       | No   | No       | No   | No     | No     | No     |
| Record Keeping | Yes                       | Yes  | Yes      | Yes  | Yes    | Yes    | Yes    |
| Group Thinking | Yes                       | Yes  | Yes/No   | No   | No     | No     | No     |
| Anonymity      | No                        | Yes  | Yes      | Yes  | Yes    | Yes    | Yes    |
| Time           | Medium                    | High | High/Low | Low  | Low    | Low    | Low    |
| Precision      | Low                       | Low  | Low      | High | High   | High   | High   |
| Accuracy       | High                      | High | High     | Low  | Medium | Medium | Medium |

## 4 Experience Using Consensus

In our work with a government regulator, we used GRL to model goals and indicators for regulations in order to support inspection-based measurement of compliance and performance [6, 24]. We needed to insure high accuracy while minimizing time spent in meetings, so we selected RTD&C. As anonymity was not a concern, we encouraged discussions in order to learn from domain experts. This work was performed on two separate sets of regulations with two different sets of clients.

For the first clients, the objective was to convert a set of *management-based regulations* to outcome-based regulations and derived goal models. Management-based regulations direct regulated organizations to engage in a planning process that targets the achievement of public goals, while offering regulated parties flexibility as to how they achieve these goals [8]. In our RTD&C meetings, each regulation was displayed on screen and read out loud. Participants then discussed what indicator was needed together with evaluation questions that would reliably measure whether the desired outcome had been achieved. How to score different levels of compliance was also discussed. The meetings' progress was initially very slow but the speed picked up as participants got comfortable with the process.

For the second clients, we decided to use RTD&C again to derive inspection/evaluation questions from a large set of *prescriptive regulations* imposing specific actions rather than specifying desired outcomes. Prescriptive regulations are more complex because why actions exist is not always apparent. However, because the modeling team had gained experience working with the first clients, progress was much faster and smoother. Each regulation was projected and read, and the participants decided the desired outcome(s), relevant indicators, and inspection questions. For weighting the indicators, the group selected RTD&C over Delphi mainly because of lower preparation time. The assignment of relative weights was relatively faster because the regulations were dealt with in groups. In general, both experiments were successful. The rigor of the process and documentation were useful and compelling, particularly for questions asked later by people not involved in the actual process.

## 5 Experience Using AHP

The second set of experiments involved the use of GRL models to support decision making during adaptation of enterprise architectures (EA) at a different Canadian

government department [1]. The rationale was that Information Systems (IS) in the enterprise provide the information utilized by decision makers in achieving organizational objectives. The EA goal model produced showed links from IS to decisions made by decision makers, and then to business objectives, providing opportunities to trace, monitor, and address change at the architecture level. In addition, the health of each information system was monitored through the use of six quality-oriented indicators. Herein, quantification is mainly required in three places: for contribution levels, for importance values, and for indicator definitions. The EA GRL model covering the main and provincial offices had 4 diagrams, 8 actors, 40 intentional elements (12 goals, 9 softgoals, 8 tasks, and 11 resources), 30 indicators, and 102 links.

In order to determine the contribution/importance/indicator quantities, we had access to four senior and busy enterprise architects. We had little time yet we wanted accuracy. As we were not looking for a face-to-face learning opportunity, we decided to use a virtual approach based on AHP. Given that for any element in the GRL model, we only had up to six contributions, pairwise comparison was deemed to be feasible;  $n(n-1)/2$  comparisons are needed for  $n$  elements to compare, so there was a maximum of 15 questions to ask per element. Questionnaires targeting the required quantities were administered to the senior architects. The data obtained was analyzed using pairwise comparison to obtain values, normalized over a 0-100 scale in the model.

The GRL model was used to assess various adaptation scenarios of the enterprise architecture. The four architects evaluated the approach after the project through questionnaires, with positive feedback. Only one criterion pertained to the quantification itself (business-IT alignment), with a good result (4.0/5). The questionnaires were also seen as quick and easy to use, and the resulting quantities were reasonable.

## 6 Lessons Learned

Based on formal and informal discussions and observations on our GRL-based quantification approaches, we learned the following:

- Quantification of GRL goal models is practical, and many approaches can be used.
- Government departments are facing increasing demands for numerical values and quantities to support program decisions. Current approaches, while useful, do not provide the necessary rigor, whereas our GRL-based approach with indicators helps accommodate the need for objectivity and precision.
- In the first department, RTD&C enabled very effective knowledge transfer from the subject matter experts to the team. An unintended side-effect was hence an improved understanding of regulations and models in both groups.
- Even for large EA models, AHP and pairwise comparisons are feasible because local decisions only require a few (often 2 or 3) elements to compare.
- The preparation of slides in the first approach took time, but the discussions that took place among the facilitators brought out some misunderstandings that needed to be clarified. Since then, we have partially automated the creation of such views.

## 7 Conclusions and Future Work

In this paper, we demonstrated the feasibility and effectiveness of quantification approaches for goal models based on experience gained at two departments of the Government of Canada. We focused on the use of relative weights with consensus in regulatory compliance and on the use of AHP in enterprise architectures. We also introduced the use of AHP to compute contribution and importance levels in GRL models for the first time. Feedback was positive in both places, with many lessons learned.

There are obvious limitations to our results. First, we collected data as we were developing and experimenting with the group decision approaches and with the modeling styles themselves. These two aspects should be better separated in the future. Second, we used two approaches that appeared to fit the tasks, however, this does not mean that they are optimal. Our results merely suggest that quantification of goal models is feasible in the policy/regulation contexts explored.

Liaskos et al. [15] indicate further research questions related to quantification that should be considered in the future. In addition, there is a need to consider confidence and uncertainty in the quantities that are “agreed” on.

**Acknowledgements.** This work was supported in part by NSERC’s Business Intelligence Network. We also thank the many collaborators and visionary people at the Government of Canada for their participation and support.

## References

1. Akhigbe, O.S.: Business Intelligence - Enabled Adaptive Enterprise Architecture. M.Sc. thesis, Systems Science, University of Ottawa (2014), <http://hdl.handle.net/10393/31012>
2. Ahn, B.S.: Compatible weighting method with rank order centroid: Maximum entropy ordered weighted averaging approach. *EJOR* 212(3), 552–559 (2011)
3. Amyot, D., Ghanavati, S., Horkoff, J., Mussbacher, G., Peyton, L., Yu, E.: Evaluating Goal Models within the Goal-oriented Requirement Language. *International Journal of Intelligent Systems* 25(8), 841–877 (2010)
4. Amyot, D., Mussbacher, G.: User Requirements Notation: The First Ten Years, The Next Ten Years. *Journal of Software* 6(5), 747–768 (2011)
5. Amyot, D., et al.: Towards Advanced Goal Model Analysis with jUCMNav. In: Castano, S., Vassiliadis, P., Lakshmanan, L.V.S., Lee, M.L. (eds.) *ER 2012 Workshops*. LNCS, vol. 7518, pp. 201–210. Springer, Heidelberg (2012)
6. Badreddin, O., et al.: Regulation-Based Dimensional Modeling for Regulatory Intelligence. In: *RELAW 2013*, pp. 1–10. IEEE CS (2013)
7. Bresciani, P., Perini, A., Giorgini, P., Giunchiglia, F., Mylopoulos, J.: Tropos: An Agent-Oriented Software Development Methodology. *Autonomous Agents and Multi-Agent Systems* 8(3), 203–236 (2004)
8. Coglianesi, C., Lazer, D.: Management - Based Regulation: Prescribing Private Management to Achieve Public Goals. *Law & Society Review* 37(4), 691–730 (2003)

9. Hassine, J., Amyot, D.: GRL Model Validation: A Statistical Approach. In: Haugen, Ø., Reed, R., Gotzhein, R. (eds.) SAM 2012. LNCS, vol. 7744, pp. 212–228. Springer, Heidelberg (2013)
10. Horkoff, J., Barone, D., Jiang, L., Yu, E., Amyot, D., Borgida, A., Mylopoulos, J.: Strategic Business Modeling: Representation and Reasoning. *Software & Systems Modeling* 13(3), 1015–1041 (2012)
11. International Telecommunication Union, Recommendation Z.151 (10/12), User Requirements Notation (URN) – Language Definition. Geneva, Switzerland (2012)
12. Jureta, I., Faulkner, S., Schobbens, P.-Y.: Clear justification of modeling decisions for goal-oriented requirements engineering. *Requirement Engineering* 13(2), 87–115 (2008)
13. Kassab, M.: An integrated approach of AHP and NFRs framework. In: Seventh Int. Conf. on Research Challenges in Information Science (RCIS), pp. 1–8. IEEE CS (2013)
14. Letier, E., van Lamsweerde, A.: Reasoning about partial goal satisfaction for requirements and design engineering. *Software Engineering Notes* 29(6), 53–62 (2004)
15. Liaskos, S., Hamidi, S., Jalman, R.: Qualitative vs. Quantitative Contribution Labels in Goal Models: Setting an Experimental Agenda. In: iStar 2013. CEUR-WS, Vol. 978, pp. 37–42 (2013)
16. Liaskos, S., Jalman, R., Aranda, J.: On eliciting contribution measures in goal models. In: 20th Int. Requirements Engineering Conference (RE), pp. 221–230. IEEE CS (2012)
17. Lilja, K.K., Laakso, K., Palomki, J.: Using the Delphi method. In: Technology Management in the Energy Smart World (PICMET), pp. 1–10. IEEE CS (2011)
18. Linstone, H.A., Turoff, M.: The Delphi method. Addison-Wesley (1975)
19. Maté, A., Trujillo, J., Franch, X.: Adding semantic modules to improve goal-oriented analysis of data warehouses using I-star. *JSS* 88, 102–111 (2014)
20. Moody, D.L., Heymans, P., Matulevičius, R.: Visual syntax does matter: improving the cognitive effectiveness of the i\* visual notation. *Req. Eng.* 15(2), 141–175 (2010)
21. Munro, S., Liaskos, S., Aranda, J.: The Mysteries of Goal Decomposition. In: iStar 2011. CEUR-WS, vol. 766, pp. 49–54 (2011)
22. Mussbacher, G., Amyot, D., Heymans, P.: Eight Deadly Sins of GRL. In: iStar 2011. CEUR-WS, vol. 766, pp. 2–7 (2011)
23. Saaty, T.L.: A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology* 15(3), 234–281 (1977)
24. Tawhid, R., et al.: Towards Outcome-Based Regulatory Compliance in Aviation Security. In: 20th Int. Requirements Engineering Conference (RE), pp. 267–272. IEEE CS (2012)
25. van Lamsweerde, A.: Requirements Engineering: From System Goals to UML Models to Software Specifications. Wiley (2009)
26. Velasquez, M., Hester, P.T.: An Analysis of Multi-Criteria Decision Making Methods. *International Journal of Operations Research* 10(2), 56–66 (2013)
27. Vinay, S., Aithal, S., Sudhakar, G.: A Quantitative Approach Using Goal-Oriented Requirements Engineering Methodology and Analytic Hierarchy Process in Selecting the Best Alternative. In: Aswatha Kumar, M., et al. (eds.) Proceedings of ICAdC. AISC, vol. 174, pp. 441–454. Springer, Heidelberg (2012)
28. Yu, E.: Towards Modelling and Reasoning Support for Early-Phase Requirements Engineering. In: 3rd Int. Symp. on Requirements Engineering, pp. 226–235. IEEE CS (1997)