Troy Costa Kohwalter

<TÍTULO DO TRABALHO>

Escolher um item. apresentada ao Programa de Pós-Graduação em Computação da Universidade Federal Fluminense, como requisito parcial para obtenção do Grau de Escolher um item.. Área de Concentração: Escolher um item..

Advisors: Prof. Dr. Esteban G. W. Clua

Prof. Dr. Leonardo G. P. Murta

Niterói

2013

Ficha Catalográfica – Esta página deve ser removida na versão a ser entregue para a banca, mas deve ser reinserida na versão final, com a ficha catalográfica fornecida pela biblioteca. Informações sobre este processo devem ser obtidas na secretaria da pós-graduação.

Clique aqui para digitar texto.

Clique aqui para digitar texto.

Escolher um item. apresentada ao Programa de Pós-Graduação em Computação da Universidade Federal Fluminense, como requisito parcial para obtenção do Grau de Escolher um item.. Área de Concentração: Escolher um item..

Aprovada em <MES> de <ANO>.

BANCA EXAMINADORA

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Prof. Dr. <NOME DO ORIENTADOR> – Orientador

UFF

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Prof. Dr. <NOME DO AVALIADOR>

<INSTITUIÇÃO DO AVALIADOR>

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Prof. Dr. <NOME DO AVALIADOR>

<INSTITUIÇÃO DO AVALIADOR>

Niterói

<ANO>

"Dedicatória(s): Elemento opcional onde o autor presta homenagem ou dedica seu trabalho" (ABNT, 2005).

**Agradecimentos**

"Elemento opcional, colocado após a dedicatória" (ABNT, 2005).

"Epígrafe: Folha onde o autor apresenta uma citação, seguida de indicação de autoria, relacionada com a matéria tratada no corpo do trabalho." (ABNT, 2005).

**Resumo**

"Elemento obrigatório, constituído de uma sequência de frases concisas e objetivas e não de uma simples enumeração de tópicos, não ultrapassando 500 palavras" (ABNT, 2005).

Palavras-chave: "Palavras representativas do conteúdo do trabalho, isto é, palavras-chave e/ou descritores, conforme a ABNT NBR 6028" (ABNT, 2005).

**Abstract**

"Elemento obrigatório, em língua estrangeira, com as mesmas características do resumo em língua vernácula" (ABNT, 2005).

Keywords: "Palavras representativas do conteúdo do trabalho, isto é, palavras-chave e/ou descritores, na língua" (ABNT, 2005).

**Lista de Ilustrações**

[Figura 1: Exemplo de figura. 2](#_Toc288562698)

**Lista de tabelas**

[Tabela 1: Exemplo de tabela 2](#_Toc288562708)

**Lista de abreviaturas e siglas**

"Elemento opcional, que consiste na relação alfabética das abreviaturas e siglas utilizadas no texto, seguidas das palavras ou expressões correspondentes grafadas por extenso. Recomenda-se a elaboração de lista própria para cada tipo" (ABNT, 2005).

**Lista de símbolos**

"Elemento opcional, que deve ser elaborado de acordo com a ordem apresentada no texto, com o devido significado" (ABNT, 2005).

Contents

[Chapter 1 – Introduction 15](#_Toc352687049)

[1.1 Motivation 15](#_Toc352687050)

[1.2 Goals 15](#_Toc352687051)

[1.3 Research Questions 15](#_Toc352687052)

[1.4 Contributions 15](#_Toc352687053)

[1.5 Organization 15](#_Toc352687054)

[Chapter 2 – Game Flow Analysis 16](#_Toc352687055)

[2.1 Introduction 16](#_Toc352687056)

[2.2 Usage 16](#_Toc352687057)

[2.3 UNDECIDED TITLE 16](#_Toc352687058)

[2.4 Related Work 16](#_Toc352687059)

[2.5 Final Considerations 16](#_Toc352687060)

[Chapter 3 – Provenance 17](#_Toc352687061)

[3.1 Introduction 17](#_Toc352687062)

[3.2 Open Provenance Model 18](#_Toc352687063)

[3.3 PROV 23](#_Toc352687064)

[3.4 Comparison Between Models 23](#_Toc352687065)

[3.5 Final Considerations 23](#_Toc352687066)

[Chapter 4 – Provenance in Games 24](#_Toc352687067)

[4.1 Introcutions 24](#_Toc352687068)

[4.2 Data Model 25](#_Toc352687069)

[4.3 Data Structure 26](#_Toc352687070)

[4.4 Information Storage 27](#_Toc352687071)

[4.5 Provenance Visualization 27](#_Toc352687072)

[4.6 Final Considerations 27](#_Toc352687073)

[Chapter 5 – Implementation 28](#_Toc352687074)

[5.1 Introduction 28](#_Toc352687075)

[5.2 SDM 28](#_Toc352687076)

[5.3 Guiding Example 28](#_Toc352687077)

[5.4 Provenance Gathering 28](#_Toc352687078)

[5.5 Provenance Analysis 28](#_Toc352687079)

[5.5.1 Node Representations 28](#_Toc352687080)

[5.5.2 Edge Representations 28](#_Toc352687081)

[5.5.3 Node Filters 28](#_Toc352687082)

[5.5.4 Edge Filters 28](#_Toc352687083)

[5.6 Final Considerations 28](#_Toc352687084)

[Chapter 6 – Evaluation 29](#_Toc352687085)

[6.1 Introduction 29](#_Toc352687086)

[6.2 Experiment Planning 29](#_Toc352687087)

[6.3 Experiment Execution 29](#_Toc352687088)

[6.4 Statistical Analysis 29](#_Toc352687089)

[6.5 Threats to Validity 29](#_Toc352687090)

[6.6 Final Considerations 29](#_Toc352687091)

[Chapter 7 – Conclusion 30](#_Toc352687092)

[7.1 Contributions 30](#_Toc352687093)

[7.2 Limitations 30](#_Toc352687094)

[7.3 Future Work 30](#_Toc352687095)

[References 31](#_Toc352687096)

[Glossary 33](#_Toc352687097)

[Appendix A – Título do Apêndice 34](#_Toc352687098)

[Annex A – Título do Anexo 35](#_Toc352687099)

[Index 36](#_Toc352687100)

# – Introduction

## Motivation

Exemplo de texto de uma dissertação ou tese. Esse exemplo mostra como referenciar figuras e tabelas. A consiste um exemplo de figura, e esta frase faz uma citação com referência cruzada a essa figura.

C:\Program Files (x86)\Microsoft Office\MEDIA\CAGCAT10\j0217698.wmf

Figura 1: Exemplo de figura.

Este parágrafo tem como propósito exibir uma citação de tabela, também utilizando referência cruzada. A Tabela 1 consiste em um exemplo de tabela.

Tabela 1: Exemplo de tabela

|  |  |  |  |
| --- | --- | --- | --- |
|  | Coluna 1 | Coluna 2 | Coluna 3 |
| Linha 1 |  |  |  |
| Linha 2 |  |  |  |

As referências bibliográficas utilizadas nesse modelo foram geradas pelo Zotero (GEORGE MASON UNIVERSITY, 2011) utilizando o estilo ABNT fornecido por ROCHA (2011).

## **Goals**

## **Research Questions**

## Contributions

## Organization

# – Game Flow Analysis

## Introduction

Exemplo de texto de uma dissertação ou tese. Esse exemplo mostra como referenciar figuras e tabelas. A consiste um exemplo de figura, e esta frase faz uma citação com referência cruzada a essa figura.

## **Usage**

## **UNDECIDED TITLE**

## Related Work

## Final Considerations

# – Provenance

## Introduction

Results of scientific experiments cannot be understood without the knowledge of the meaning of data and circumstances occurred during their creation. This type of knowledge includes data provenance (DAVIDSON; FREIRE, 2008; FREIRE *et al.*, 2008). The provenance term is well understood in the context of art or digital libraries, where historical documentation refers to an object’s life cycle (PREMIS WORKING GROUP, 2005). Data provenance in scientific experimentation has become such an important topic that workshops and conferences on the subject were specifically created (SIMMHAN; PLALE; GANNON, 2005).

The *International Provenance and Annotation Workshop* (IPAW) (MOREAU *et al.*, 2002) was one of the first data provenance workshops to be created. In each edition, the scientific community listed challenges of data provenance to be solved and receives many scientists work with suggested solutions. During IPAW’06, participants were interested in questions about provenance for the use in digital data, involving topics related to documentation, data annotation and data derivations (BOSE; FOSTER; MOREAU, 2006). As a result, the first model of digital provenance, the *Open Provenance Model* (OPM) (MOREAU *et al.*, 2007), was created. The OPM has been designed to address the issues raised during the *Provenance Challenge* (“Provenance Challenge WIKI”, 2010).

Later, another provenance model, PROV (GIL; MILES, 2010), was developed by the provenance incubator group W3C (GIL *et al.*, 2009). According to the group, provenance of digital objects represents the object’s origins and PROV is a proposed specification to represent these provenance records. These records contain descriptions of the entities and activities involved in producing and delivering or otherwise influencing a given object. The usage of provenance, regardless of the model, provides a critical foundation for assessing the authenticity of data, enabling reliability and reproducibility and is crucial component of workflow systems (GIL *et al.*, 2007; GROTH; MOREAU, 2010).

When PROV was proposed, the OPM model was already being used in several approaches. However, the fact that PROV is supported by the W3C makes the possibility becoming the default provenance model, making the migration from OPM to PROV a possibility in the near future. With this, the aim of this chapter is to present a study of the digital provenance models, as well as comparing those models, pointing out their similarities and differences.

As such, this chapter is organized as follow: …

## **Open Provenance Model**

The *Open Provenance Model* emerged as a result from the *Provenance Challenges* proposed in the context of IPAW. The *Provenance Challenges* came in four editions, one for each year from 2006 to 2010 and OPM resulted from the first two challenges and was used on the third challenge:

*1st Challenge*: Aimed to provide a forum for the community to understand the capabilities of different provenance systems and express their provenance representations.

*2nd Challenge*: Aimed to establish interoperability between systems through exchange of provenance information.

*3rd Challenge*: Evaluate the OPM practically, from an inter-operability view-point.

The *Open Provenance Model* is a provenance model designed to meet the following requirements (MOREAU *et al.*, 2007):

* To allow provenance information to be exchanged between systems, by means of compatibility layer based on a shared provenance model.
* To allow developers to build and share tools that operates on such a provenance model.
* To define provenance in a precise, technology-agnostic manner.
* To support a digital representation of provenance for any “thing”, whether produced by computer systems or not.
* To allow multiple levels of descriptions to coexist.
* To define a core set of rules that identify the valid inferences that can be made on provenance representation.

In Open Provenance Model, it is assumed that provenance of objects is represented by an annotated causality graph, which is a directed acyclic graph enriched with annotations capturing further information pertaining to execution. According to MOREAU *et al.* (2007), a provenance graph is a record of a past or current execution, and not a description of something that could happen in the future.

The causality graph is composed of nodes that can represent *Artifacts*, *Processes* and *Agents*. *Artifacts* are an immutable piece of state that can represent a physical object or a digital representation in a computer system. *Processes* are actions or a sequence of actions performed or caused by artifacts and results in new artifacts. *Agents* are contextual entities acting as a catalyst of a process that can enable, facilitate, control or affect its execution. The edges of the graph belong to one of the categories described in , representing a causal dependency between its source, denoting the effect, and its destination that denotes the cause. Below are some important definitions in the Open Provenance Model according to MOREAU *et al.* (2007).

# 

Figure 1: Edges and Usage of Timestamps in OPM. Source: [Moreau et al. 2011].

**Causal Relationship**: Represented by an arc and denotes the presence of a causal dependency between the source (effect) and the destination (cause).

**Artifact Used by a Process**: A [*used*] edge from process to an artifact is a causal relationship intended to indicate that the process required the availability of the artifact to be able to complete its execution. When several artifacts are connected to a same process by multiple [*used*] edges, all of them were required for the process to complete.

**Artifacts Generated by Processes**: A [*was generated by*] edge from an artifact to a process is a causal relationship intended to mean that the process was required to initiate its execution in order to generate the artifact. When several artifacts are connected to the same process by multiple [*was generated by*] edges, the process must begin for all of them to be generated.

**Process Triggered by Process**: An edge [*was triggered by*] from a process P2 to a process P1 is a causal dependency that indicates that the start of process P1 was required for P2 to be able to complete.

**Artifact Derived from Artifact**: An edge [*was derived from*] from artifact A2 to artifact A1 is a causal relationship that indicates that artifact A1 should have been generated for A2 to be generated. The piece of state associated with A2 is dependent on the presence of A1 or on the piece of state associated with A1.

**Process Controlled by Agent**: An edge [*was controlled by*] from a process P to an agent Ag is a causal dependency that indicates that agent Ag controlled the start and end of process P.

**Role**: Designates an artifact or agent's function in a process.

In , the edge [*used*] say that a process used an artifact, while the [*was generated by*] edge an artifact was generated by a process. The letter "R" represents the roles under which these artifacts were used since a process may have used several artifacts. Likewise, many artifacts may have been generated by a process, and each would have a specific role. Roles are only meaningful in the context of the process where they are defined, and they are not defined by the OPM itself, but by the application domains. Roles are used on OPM just to distinguish the involvement of artifacts in processes.

The edge [*was controlled by*] means the process was caused by an agent, essentially acting as a catalyst or controller. Since a process may have been controlled by several agents, their roles are also identified as controllers. This type of dependency represents a control relationship and not a data derivation. The edge [*derived from*] assert that artifact A2 was derived from another artifact A1, giving an oriented dataflow view of the provenance. In contrast to the edge [*was derived from*], an edge [*was triggered by*] allows a process to have an oriented view of past executions.

Moreover, the Open Provenance Model allows causality graphs to be used with time information. In this model, time is not used for deriving causality, but to validate causality claims, since if the same time clock is used to measure the time for both the effect and cause, then the time of an effect should be greater than the time of its cause.

In addition, time may be associated to *instantaneous occurrences* in a process. There are four types of this occurrences, being denoted as *creation* and *use* for artifacts and *starting* and *ending* for processes. Given that time may be observed by someone, its accuracy is limited by the clock and the notion of time. This way, the model allows for an interval of accuracy to support the granularity used to represent time. With this, it is possible to state that an artifact was used no earlier than time t1 and no later than time t2, as an example. This rationale is analogous for processes.

indicates how time information can be expressed in the model. For [*used*] and [*was generated by*] edges, one timestamp can be used to express when the event happened. For [*was controlled by*] edge two timestamps marks when the process started and terminated. For [*was derived from*] and [*was triggered by*] edges, one timestamp to indicate when the artifact was used. Despite using timestamp, the time of occurrence itself is not enough to imply causality. The fact that process P1 happened before P2 is not enough information to infer that P1 caused P2 to happen.

Finally, the Open Provenance Model has defined the notion of a graph based on a set of syntactic rules and topological constraints. The provenance graph captures causal dependencies that can be summarized by means of transitive closure. Because of this, a set of completion rules and inferences can be used in the graph.

For completion rules, there is the artifact elimination, also known as forward transformation. shows such transformation. The edge [*was triggered by*] can be obtained from the existence of [*used*] and [*was generated by*] edges. Also in the same figure, there is another completion rule, called artifact introduction, which establishes that the [*was triggered by*] edge is hiding the existence of an artifact used by P2 and generated by P1. The completion rules allow the establishment of the existence of some artifacts but it does not make explicit their identities. This is the consequence of using [*was triggered by*], which is a composition of [*used*] and [*was generated by*]. On the other hand, presents a completion rule regarding *process introduction*. The edge [*was derived from*] hide the presence of an intermediary process. However, the converse rule does not work without some internal knowledge of P, which is fundamental to ascertain if there is an actual dependency between A1 and A2.



Figure 2: Artifact introduction and elimination. Source: [Moreau et al. 2011].

When users want to find out the causes of an artifact or a process, their interest is in indirect causes that involve multiple transitions. For this purpose, a set of new relationships was created:



**Figure 3: Process introduction. Source:** (MOREAU *et al.*, 2007)**.**

**Multi-step "wasDerivedFrom"**: An artifact *a1* was derived from *A2* (possibly using multiple steps), written as *a1🡪\* a2*, if *a1* was derived from *a2* or from an artifact that was itself derived from *a2* (possibly using multiple steps). In other words, it is the transitive closure of the edge [*was derived from*]. It expresses that artifact *a2* had an influence on artifact *a1.*

**Secondary Multi-Step Edges**:

**Process *p* used artifact *a* (possibly using multiple steps)**:written *p 🡪\* a*, if *p* used an artifact *a* or an artifact that derived *a* (possibly using multiple steps).

**Artifact *a* was generated by process *p* (possibly using multiple steps)**:written *a* 🡪\* *p*, ifa or an artifact that derived *a* (possibly using multiple steps) that was generated by *p.*

**Process *p1* was triggered by process *p2* (possibly using multiple steps)**:written *p1* 🡪\* *p2,* if *p1* used an artifact that was generated or was derived from an artifact (possibly using multiple steps) that was itself generated by *p2.*

Multi-step edges can be inferred from single step edges by eliminating artifacts that occur in chains of dependencies. Analyzing , it is possible to infer that process *p2* was triggered by *p1*, omitting the fact that *p2* used *a3*, which was derived from *a2* that in turn was derived from *a1*, which was generated by *p1*. Other inferences are also illustrated in .



Figure 4: Inference. Source: [Moreau et al. 2011].

Lastly, the *Open Provenance Model* has a modular design as illustrated by Figure 5. However, specifications for all layers in the design have not been produced yet. At the bottom layer is located the abstract model (MOREAU *et al.*, 2007). On the left side, a serialization to *xml*, defined by OPMX (The Open Provenance Model XML Schema) (MOREAU; GROTH; *et al.*, 2010), and a mapping to RDF with OPMV (The Open Provenance Model Vocabulary) (ZHAO, 2010) and OPMO (The Open Provenance Model OWL Ontology) (MOREAU; DING; *et al.*, 2010). Those are the only specifications produced, along with the *Open Provenance Model Java Library* (MOREAU, 2010b), a JAXB-generated Java Library used by *OPM Toolbox* (MOREAU, 2010a) for creating a Java representation of OPM graphs and serializing them to or from a *xml*  file. With the development of PROV, these other OPM specifications were left unfinished.

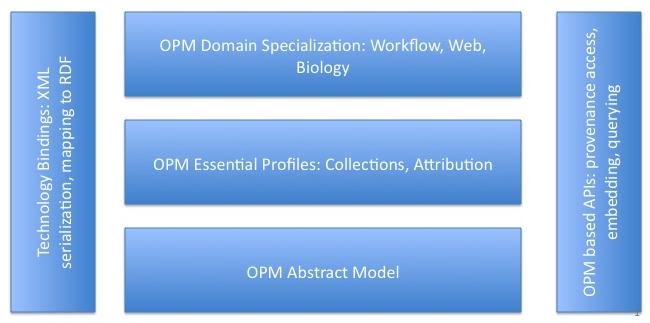


Figure 5: OPM Layered Architecture

## **PROV**

## Comparison Between Models

## Final Considerations

# – Provenance in Games

## Introcutions

In this work we propose the adoption of provenance in the context of games. For this, it is necessary to map each node of a provenance graph to elements that can be represented in the game. As was mentioned earlier, the Open Provenance Model has three types of nodes: *Artifacts*, *Process* and *Agents*. In order to map them, it is necessary to find similarities in a game context.

Starting with *Artifacts*, their provenance definition states that they are "*an immutable piece of state that can represent a physical object* […]". Its definition already gives a clue on which role they can represent in the game context: objects. An object can be anything used in the game, for example in the case of an RPG, *Artifacts* can represent weapons, potions, legendary artifacts, magical objects, etc. It can represent anything meaningful to the development of the game history.

On the other hand, *agents* "*are contextual entities acting as a catalyst of a process that can enable, facilitate, control or affect its execution*". In a game context, agents can be mapped as people represented in the game, non-playable characters (NPCs), monsters, and players.

Lastly, *Processes* according to its definition are "*actions or a sequence of actions performed or caused by artifacts* […]". So, in a game context, *Processes* can be viewed as actions or events made by living or intelligent entities that are present in the game. Note that it was made a difference between living and intelligent. This difference is important to mention because, for example, in an RPG environment a sword can be expressed as an agent because this sword has intelligence on its own. Despite being an object (sword), it can think and by an extent act, therefore it cannot be considered only as an object. It can also be as complex as being both an object and an agent at the same time.

Now, with all three types of nodes mapped into the game context, it is also necessary to map their causal relations to create the provenance graph. The Open Provenance Model defines a few causal relations which can be used similarly to their original context, but can be extended to be more suitable to the game context if necessary. Also, the Open Provenance Model can deal well with the aspect of time, which can be heavily explored in games, especially on games focused on storytelling, recording when each event happened and using this information to generate other events.

To generate actions and control events, each NPC in the game will require a decision tree in order to control his actions, providing an array of behavior possibilities. Event triggers can also be controlled by decisions tree. The next subsection describes which information is stored in actions, events, objects, and agents. We also describe how the impact decisions tree can be achieved by actions and how this information can be processed in order allow further provenance analysis.

## **Data Model**

Actions can be represented by a series of attributes that describe it and the context it was involved, allowing the creation of a provenance graph. As illustrated by , every action needs some information: a reason for its existence, why the action was performed, what triggered it, and who performed the action. In addition, the time of its occurrence can be important depending of the reason of using provenance. The main reason of using provenance, as discussed in this paper, is to produce a graph containing details that can be tracked to determine why something occurred the way it did. Therefore, with this assumption, the time of the action, the person who did it, what the action produced, and what it affect are recorded for further analysis.

Events also work in a similar way as action, with the difference in who triggered them, since events are not necessary tied to persons. For objects, its name, type, location, importance and the events that are generated by it can be stored to aid in the construction of the graph. Lastly, agents can have their names, attributes, goals, and current location recorded. illustrates this model.



Figure 6: Data model diagram. Gray classes represents provenance classes.

## **Data Structure**

In order to store all the necessary data to be used later for provenance reasons, it is required a storage structure. Depending on the information structure, it is possible to use the structure itself for inference in provenance, simplifying some unnecessary information.

Considering the generation of actions, which are executed by an entity, the action information can be stored in a list. Each entity will then have a list of actions that contains all executed actions. This allows inferring who executed each action by simply looking at whose list it belongs to, without the need to explicitly say who executed the action. For event analysis it is possible to use an analogous approach. In the case there was an external influence that resulted in the triggering of an action, then the generated action is linked to the influence, which also has links to the actions that generated the influence. Since actions belong to lists that are linked to entities, then it is possible to infer who influenced the outcome of the action by following the links.

Entities present in a scene, or place, can be represented in a similar way as actions. Each scene has a list of entities that belong to it. To represent a world, a list of scenes is created, which in turn contains list of entities that are in the scene. Each entity in turn has a list of performed actions, which have links to influences. Using this structure, it is possible to simplify some inferences in the provenance model, such as to show only relevant actions, which has external influences, to evaluate the outcome of a game session. An example of such structure is shown at , where the world has a list of scenes, each scene a list of all entities, and lastly each entity has a list of performed actions.



Figure 7: Example of structure

## Information Storage

## Provenance Visualization

The purpose of collecting information during a game session is to be able to use provenance techniques to analyze and infer the reasons of the outcome. In the previous sections, we introduced a framework to store such information. However, not all stored information is relevant for the analysis. The provenance graph contains replication of actions that did not provoke any significant change. These elements act as noise and can be omitted during provenance analysis by using completion and inference rules.

With the aim of finding actions that had an impact in the story, the actions that did not cause any dramatic change are omitted using multi-step inference rules. As an example, we may have a player in combat with an enemy and only after a few rounds it falls under the player's attacks. With the proposed framework, every round creates a node to represent the action taken by the player, which is attacking the enemy. This causes replication of data that is unnecessary for analysis, so it is possible to reduce all these individual attack nodes to simply one node.

However, that is not always true. The player could have made other actions against the enemy, which are also considered a form of attack, such as casting a spell, or a special attack maneuver, or even healing himself in order to survive. These actions are not duplicated, but can still be encapsulated for a superficial analysis, and if necessary can be expanded for a detailed analysis. Note that all collected information is preserved and the only change made is on how it is displayed. Since provenance is an analysis from the present to the past, the outcome of the battle is already known and can be used to decide which actions were relevant. If the player was victorious with minor challenge, did not suffer severe wounds, or barely used any resources at his disposal, then the entire combat can be simplified to just one node representing that the player attacked the enemy and was victorious. However, if the combat was challenging or the player lost, it is interesting to show all action nodes for analysis so the player can deeply understand the combat and decide what and when something went wrong. The ways to determine which groups of actions can be encapsulated to only one node, omitting all events in that group, is a future work of this research. However, such decisions are also dependable of the context.

## Final Considerations

# – Implementation

## Introduction

Exemplo de texto de uma dissertação ou tese. Esse exemplo mostra como referenciar figuras e tabelas. A consiste um exemplo de figura, e esta frase faz uma citação com referência cruzada a essa figura.

## **SDM**

## **Guiding Example**

## Provenance Gathering

## Provenance Analysis

### Node Representations

### Edge Representations

### Node Filters

### Edge Filters

## Final Considerations

# – Evaluation

## Introduction

Exemplo de texto de uma dissertação ou tese. Esse exemplo mostra como referenciar figuras e tabelas. A consiste um exemplo de figura, e esta frase faz uma citação com referência cruzada a essa figura.

## **Experiment Planning**

## **Experiment Execution**

## Statistical Analysis

## Threats to Validity

## Final Considerations

# – Conclusion

## Contributions

Exemplo de texto de uma dissertação ou tese. Esse exemplo mostra como referenciar figuras e tabelas. A consiste um exemplo de figura, e esta frase faz uma citação com referência cruzada a essa figura.

## **Limitations**

## **Future Work**

# References

BOSE, Rajendra; FOSTER, Ian; MOREAU, Luc. Report on the International Provenance and Annotation Workshop: (IPAW’06) 3-5 May 2006, Chicago. *SIGMOD Rec.*, v. 35, n. 3, p. 51–53, set. 2006. Accessed: 2 abr. 2013.

DAVIDSON, Susan B.; FREIRE, Juliana. Provenance and scientific workflows: challenges and opportunities. SIGMOD ’08, 2008, New York, NY, USA. *Anais*... New York, NY, USA: ACM, 2008. p. 1345–1350. Available: <http://doi.acm.org/10.1145/1376616.1376772>. Accessed: 26 mar. 2013.

FREIRE, J. *et al.* Provenance for Computational Tasks: A Survey. *Computing in Science Engineering*, v. 10, n. 3, p. 11–21, 2008.

GIL, Yolanda *et al.* Examining the Challenges of Scientific Workflows. *Computer*, v. 40, n. 12, p. 24–32, dez. 2007. Accessed: 2 abr. 2013.

GIL, Yolanda *et al.* *W3C Provenance Incubator Group*. Available: <http://www.w3.org/2005/Incubator/prov/wiki/Main\_Page>. Accessed: 22 mar. 2013.

GIL, Yolanda; MILES, Simon. *PROV Model Primer*. Available: <http://www.w3.org/TR/prov-primer/>. Accessed: 21 mar. 2013.

GROTH, Paul; MOREAU, Luc. *PROV-Overview*. Available: <http://www.w3.org/TR/prov-overview/>. Accessed: 26 mar. 2013.

MOREAU, Luc *et al.* *IPAW*. Available: <http://www.ipaw.info/>. Accessed: 2 abr. 2013.

MOREAU, Luc; DING, Li; *et al.* *Open Provenance Model (OPM) OWL Specification*. Available: <http://openprovenance.org/model/opmo>. Accessed: 2 abr. 2013.

MOREAU, Luc; GROTH, Paul; *et al.* *Open Provenance Model (OPM) XML Schema Specification*. Available: <http://openprovenance.org/model/opmx>. Accessed: 2 abr. 2013.

MOREAU, Luc. *OPM Toolbox*. Available: <http://openprovenance.org/toolbox.html>. Accessed: 2 abr. 2013a.

MOREAU, Luc. *OPM4J: The Open Provenance Model Java Library*. Available: <http://openprovenance.org/java/site/1\_1\_8/apidocs/org/openprovenance/model/package-summary.html>. Accessed: 2 abr. 2013b.

MOREAU, Luc *et al.* The Open Provenance Model core specification (v1.1). *In: Future Generation Computer Systems*, v. 27, n. 6, p. 743–756, 2007. Accessed: 10 jun. 2012.

PREMIS WORKING GROUP. *Data Dictionary for Preservation Metadata*. Technical, no Final Report of the PREMIS Working Group. Preservation Metadata: Implementation Strategies (PREMIS), 2005. Available: <http://www.oclc.org/research/projects/ pmwg/premis-final.pdf>.

*Provenance Challenge WIKI*. Available: <http://twiki.ipaw.info/bin/view/Challenge/>. Accessed: 26 mar. 2013.

SIMMHAN, Yogesh L.; PLALE, Beth; GANNON, Dennis. A survey of data provenance in e-science. *SIGMOD Rec.*, v. 34, n. 3, p. 31–36, set. 2005. Accessed: 2 abr. 2013.

ZHAO, Jun. *Open Provenance Model Vocabulary Specification*. Available: <http://open-biomed.sourceforge.net/opmv/ns.html>. Accessed: 2 abr. 2013.

# Glossary

"Elemento opcional, elaborado em ordem alfabética" (ABNT, 2005).

# Appendix A – Título do Apêndice

"Elemento opcional. O(s) apêndice(s) são identificados por letras maiúsculas consecutivas, travessão e pelos respectivos títulos. Excepcionalmente utilizam-se letras maiúsculas dobradas, na identificação, quando esgotadas as 23 letras do alfabeto" (ABNT, 2005).

# Annex A – Título do Anexo

"Elemento opcional. O(s) anexo(s) são identificados por letras maiúsculas consecutivas, travessão e pelos respectivos títulos. Excepcionalmente utilizam-se letras maiúsculas dobradas, na identificação dos anexos, quando esgotadas as 23 letras do alfabeto" (ABNT, 2005).

# Index

"Elemento opcional, elaborado conforme a ABNT NBR 6034" (ABNT, 2005).