

> Remplacer les QCM par le dialogue naturel au sein des jeux sérieux



*Paris > LIP6 > DESIR > MOCAH*

Franck.Dernoncourt@lip6.fr  
Encadrant : Jean-Marc Labat

24 Mai 2012



1. Introduction

2. État de l'art

3. Approche

4. Conclusions et perspectives

5. Références

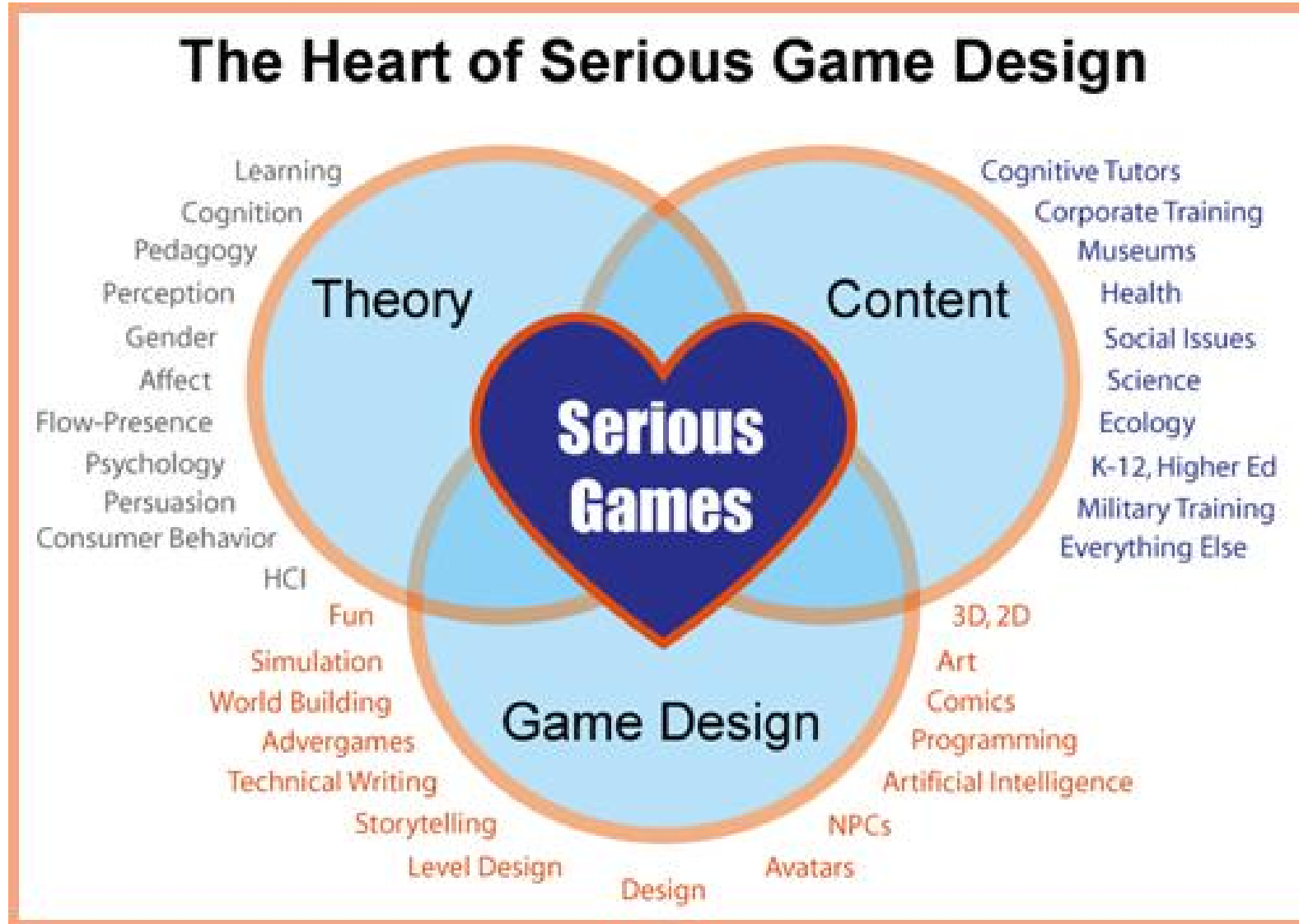
## 1) Introduction



### **Objectif :**

Conception d'un **agent conversationnel** capable de comprendre des énoncés en **langage naturel** dans un **cadre sémantique restreint** (QCM dans un jeu sérieux).

## The Heart of Serious Game Design





### Agents conversationnels :

Les agents conversationnels se divisent en deux classes principales :

- les agents conversationnels non orientés tâche,
- les agents conversationnels orientés tâche.



### **Agents conversationnels :**

Les agents conversationnels se divisent en deux classes principales :

- les agents conversationnels non orientés tâche,
- les agents conversationnels orientés tâche.

Les agents conversationnels orientés tâche sont eux-mêmes classés usuellement en deux catégories :

- les agents conversationnels orientés service,
- les agents conversationnels éducatifs.



### Agents conversationnels :

Les agents conversationnels se divisent en deux classes principales :

- les agents conversationnels non orientés tâche,
- les agents conversationnels orientés tâche.

Les agents conversationnels orientés tâche sont eux-mêmes classés usuellement en deux catégories :

- les agents conversationnels orientés service,
- **les agents conversationnels éducatifs.**

## 1) Introduction



L'existant :

QCM :

« Quelles sont les limites des algorithmes évolutionnistes ? »

- Ils ne trouveront jamais l'optimal global ;
- Il est en général impossible de faire de l'apprentissage en temps réel avec ;
- Ils ne s'appliquent pas à la plupart des problèmes d'optimisation ;
- Ils n'ont été que peu étudiés dans la littérature.



## 1) Introduction



L'existant :

QCM :

« Quelles sont les limites des algorithmes évolutionnistes ? »

- Ils ne trouveront jamais l'optimal global ;
- **Il est en général impossible de faire de l'apprentissage en temps réel avec ;**
- Ils ne s'appliquent pas à la plupart des problèmes d'optimisation ;
- Ils n'ont été que peu étudiés dans la littérature.

## 1) Introduction



### Objectif

Bot : « J'ai entendu dire que les algorithmes évolutionnistes permettraient de résoudre notre problème. Quelles sont les limites de ce genre d'algorithmes ? »

Humain : « Ces algorithmes prennent beaucoup de temps à s'exécuter et peuvent prendre beaucoup de place en mémoire. »

1. Introduction

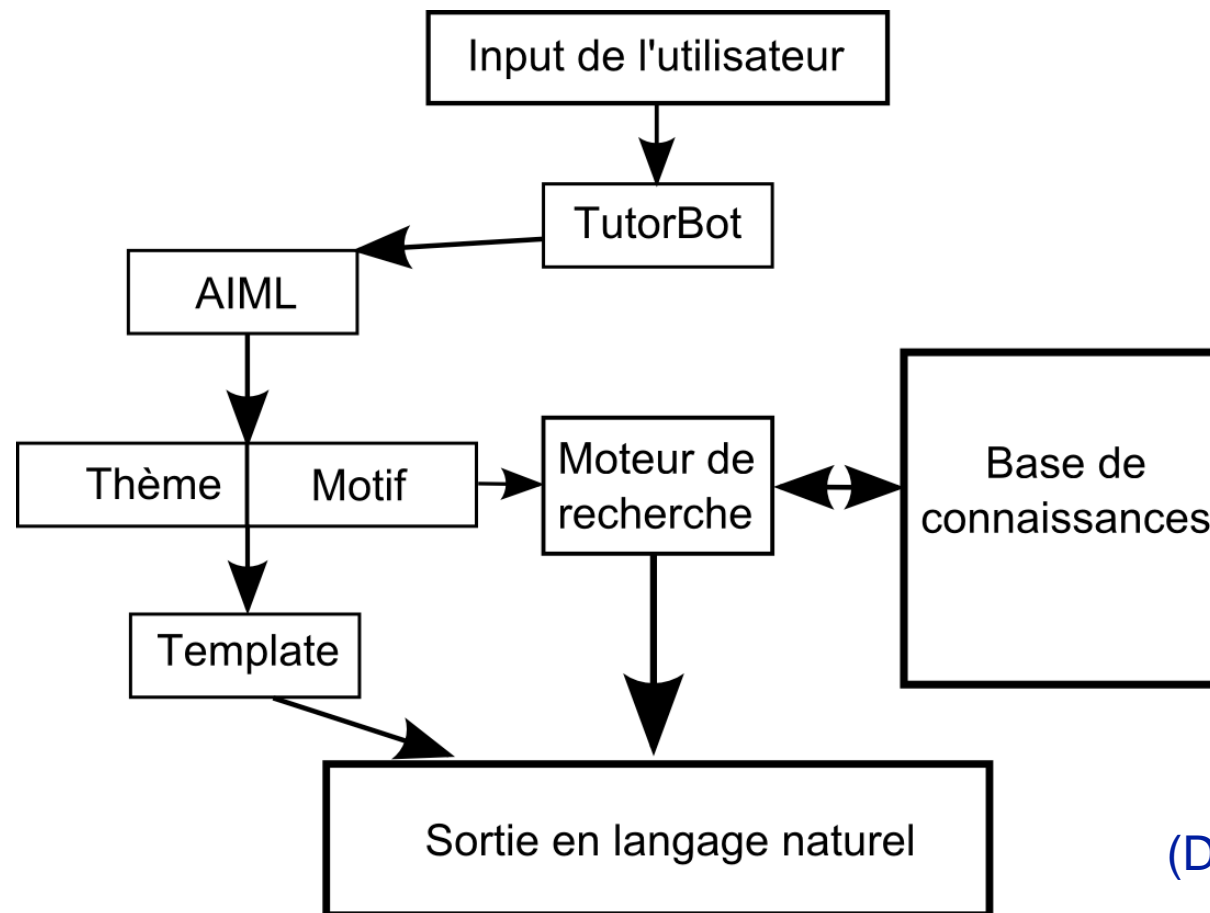
2. État de l'art

3. Approche

4. Conclusions et perspectives

5. Références

### Architecture classique :



(De Pietro, 2005)

## 2) État de l'art



### AIML :

```
<category><pattern>Do I like cheese</pattern>  
<template>How would I know if you like cheese or not?</template>  
</category>
```

## 2) État de l'art



### AIML :

```
<category><pattern>Do I like cheese</pattern>  
<template>How would I know if you like cheese or not?</template>  
</category>
```

#### Exemple 1 :

Human: Hi.

Cleverbot: How clever are you?

Human: Very clever.

Cleverbot: How clever?

## 2) État de l'art



### AIML :

```
<category><pattern>Do I like cheese</pattern>  
<template>How would I know if you like cheese or not?</template>  
</category>
```

#### Exemple 1 :

Human: Hi.

Cleverbot: How clever are you?

Human: Very clever.

Cleverbot: How clever?

#### Exemple 2 :

Human: Hi.

Cleverbot: Are you a boy?

Human: Yes I am a boys.

Cleverbot: Are you a girl?



### Systemes d'évaluation des réponses libres :

Motivés par les résultats expérimentaux montrant :

- les limites des QCM en tant qu'outil d'évaluation de la connaissance des apprenants (Whittington et Hunt, 1999)
- la complémentarité des QCM avec les réponses libres (Anbar, 1991)



## 2) État de l'art



### Systemes d'évaluation des réponses libres :

Project Essay Grade, dirigé par Ellis Batten Page à l'université Duke:

Utilisation des caractéristiques stylistiques de la réponse tels la taille des mots et le nombre de prépositions, pour prédire la note du correcteur humain (Page, 1968).

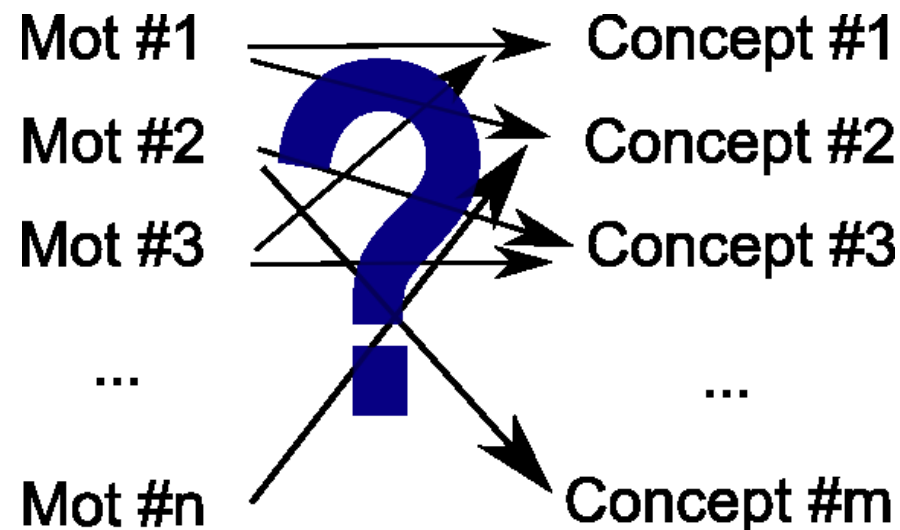
Humain : « Ces algorithmes prennent beaucoup de temps à s'exécuter et peuvent prendre beaucoup de place en mémoire. »

## 2) État de l'art



### Systemes d'évaluation des réponses libres :

LSA : Comprendre à quel concept le mot fait référence dans le contexte donné.



Humain : « Ces algorithmes prennent beaucoup de temps à s'exécuter et peuvent prendre beaucoup de place en mémoire. »

## 2) État de l'art



### Systemes d'évaluation des réponses libres :

BLEU : (méthode issue de la traduction automatique) la note donnée par BLEU au texte candidat se base sur le nombre de N-grammes communs entre le texte candidat et les textes modèles.

Humain : « Ces algorithmes prennent beaucoup de temps à s'exécuter et peuvent prendre beaucoup de place en mémoire. »

Trace 1 : « La limite principale est qu'ils mettent du temps à s'exécuter ».

...



1. Introduction

2. État de l'art

3. Approche

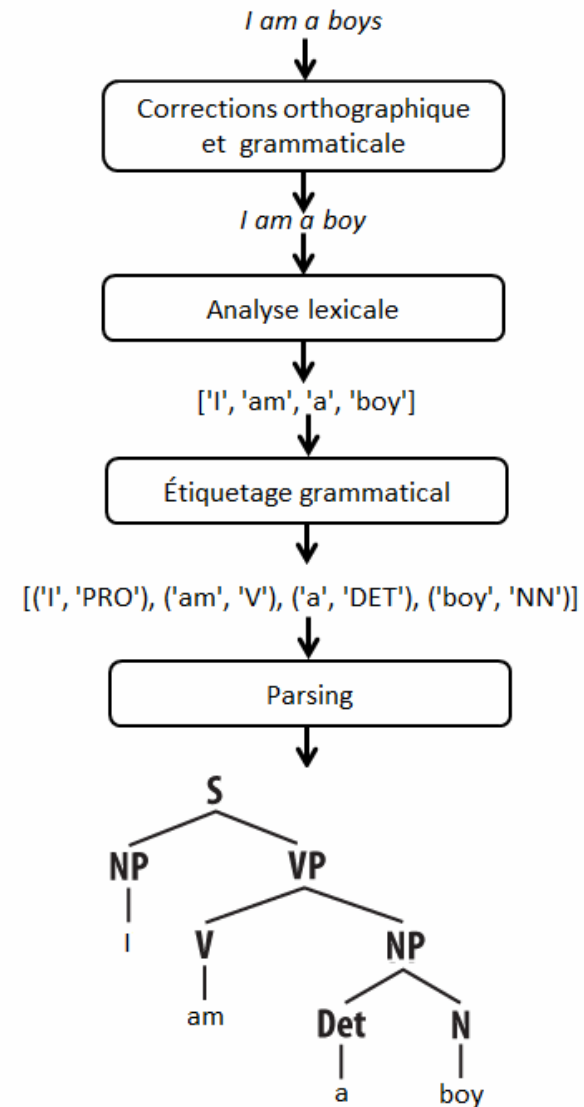
4. Conclusions et perspectives

5. Références

### 3) Approche



## Notre approche :



### 3) Approche



#### Particularités de l'approche :

- La réponse de l'apprenant n'est pas notée par rapport à des réponses modèles, mais est reliée à un QCM sous-jacent ;
- Une interaction est possible avec l'apprenant, car le système a la forme d'un agent conversationnel.

### 3) Approche



## Imprécision :

Humain : « Ces algorithmes prennent beaucoup de temps à s'exécuter et peuvent prendre beaucoup de place en mémoire. »

Bot : « Qu'entends-tu par beaucoup ? »

### 3) Approche



#### Ambiguïté structurelle :

*Groucho Marx: While hunting in Africa, I shot an elephant in my pajamas. How an elephant got into my pajamas I'll never know.*

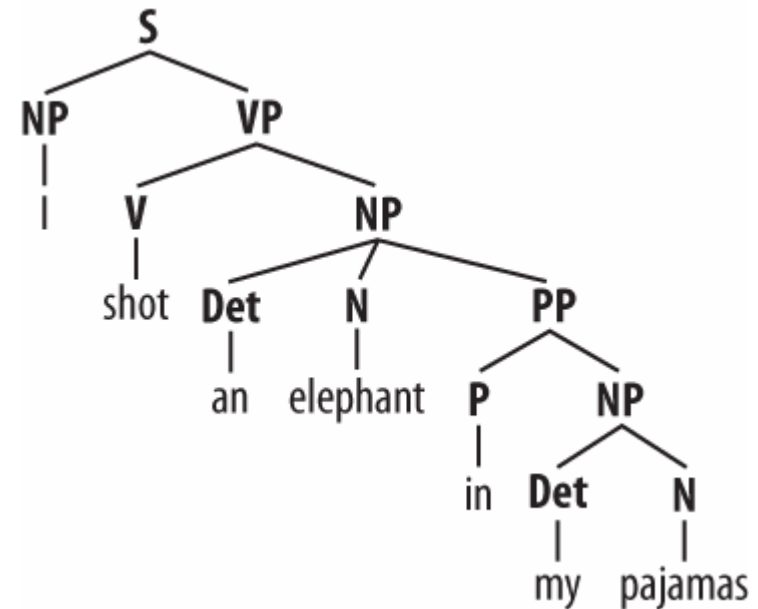
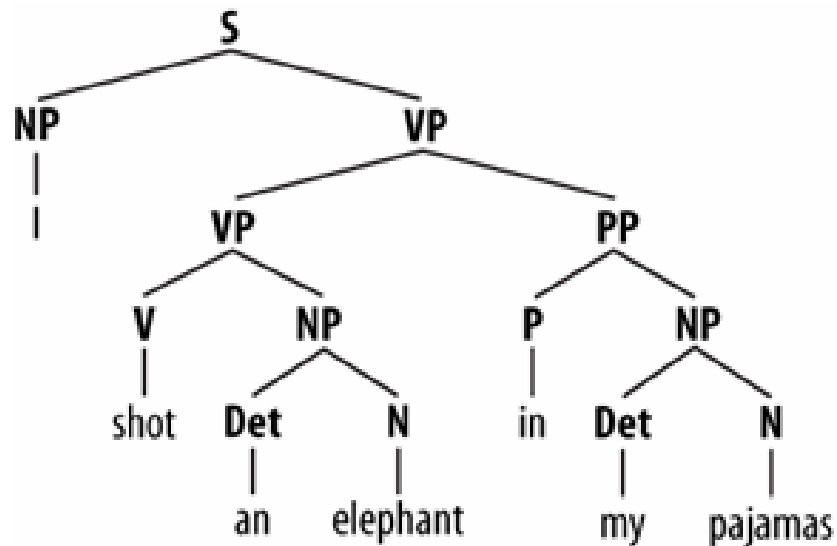


### 3) Approche



## Ambiguïté structurelle :

*Groucho Marx: While hunting in Africa, I shot an elephant in my pajamas. How an elephant got into my pajamas I'll never know.*



### 3) Approche



En résumé :

- 1) *Analyse syntaxique*
- 2) *Analyse sémantique*
- 3) *Évaluation*
- 4) *Apprentissage du système*

### 3) Approche



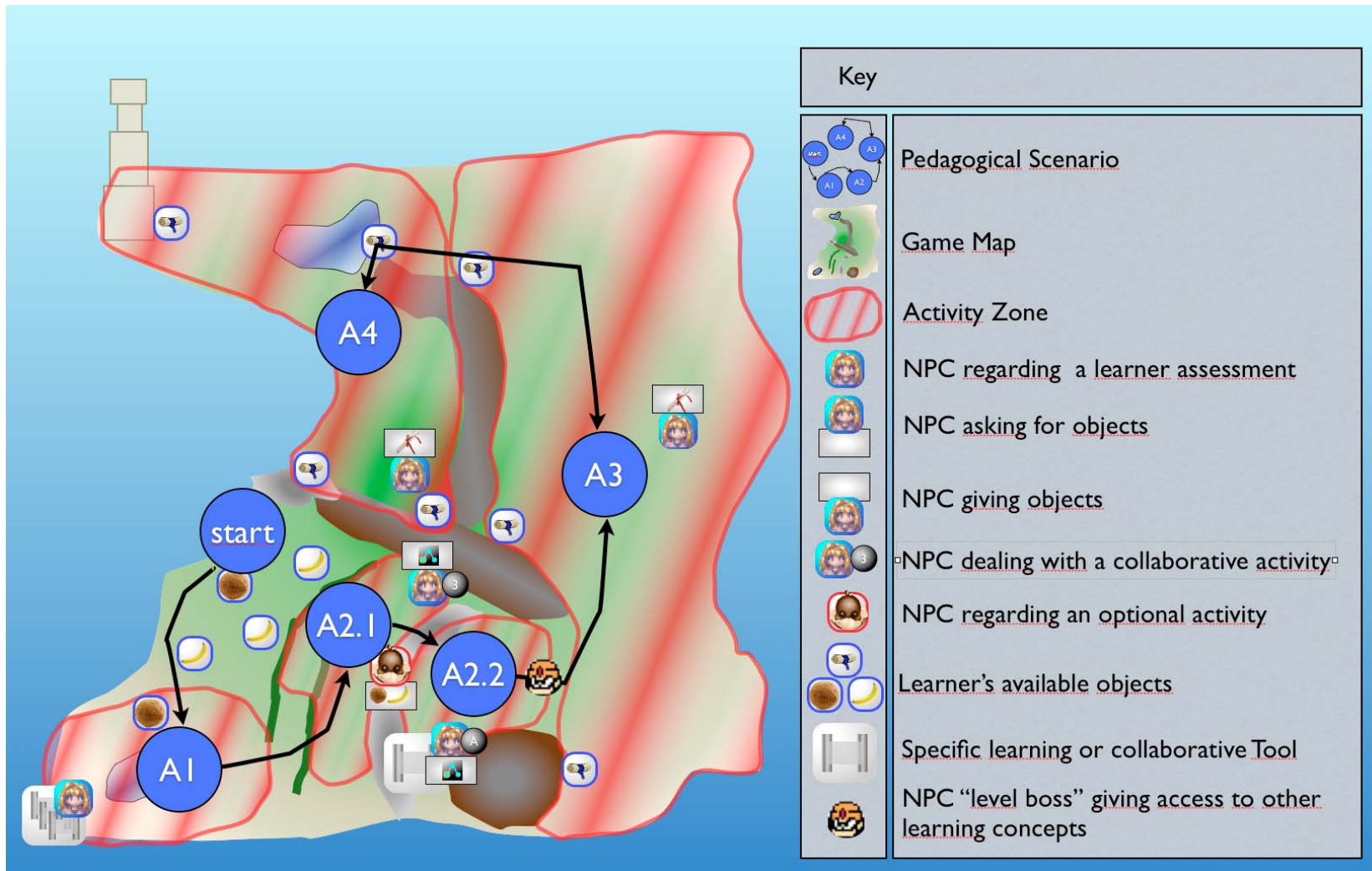
## Learning Adventure (Carron, 2010)



### 3) Approche



## Learning Adventure



### 3) Approche



## Learning Adventure

**D : dans mon travail quotidien :**

- 1. je veille à ce que mes tâches et mes objectifs soient très précisément définis +4
- 2. je ne crains pas de faire valoir mon point de vue dans les réunions +3
- 3. je peux travailler avec des personnes très différentes si leur contribution à la mission est réelle +0
- 4. je mets un point d'honneur à être au courant des idées nouvelles et à identifier les personnes nouvelles +0
- 5. je sais généralement trouver les bons arguments pour réfuter les idées ou opinions infondées +0
- 6. j'ai tendance à voir le schéma général, là où d'autres ne voient que des détails disparates +0
- 7. être très occupé me procure beaucoup de satisfaction +0
- 8. je m'intéresse réellement à la connaissance des autres +3

**E : si on me confie soudain une mission difficile, en temps limité et avec des personnes peu connues :**

- 1. je trouve que mon imagination est bridée par le travail de groupe +0
- 2. je trouve que j'ai des compétences particulièrement adaptées à ce genre de situation +0
- 3. mes sentiments ne perturbent que rarement mes raisonnements
- 4. je me bats pour construire une organisation efficace
- 5. je peux travailler avec des personnes très différentes, tant dans leurs compétences que dans leurs idées
- 6. je pense que pour faire passer ses idées, il est parfois nécessaire d'être très insistants

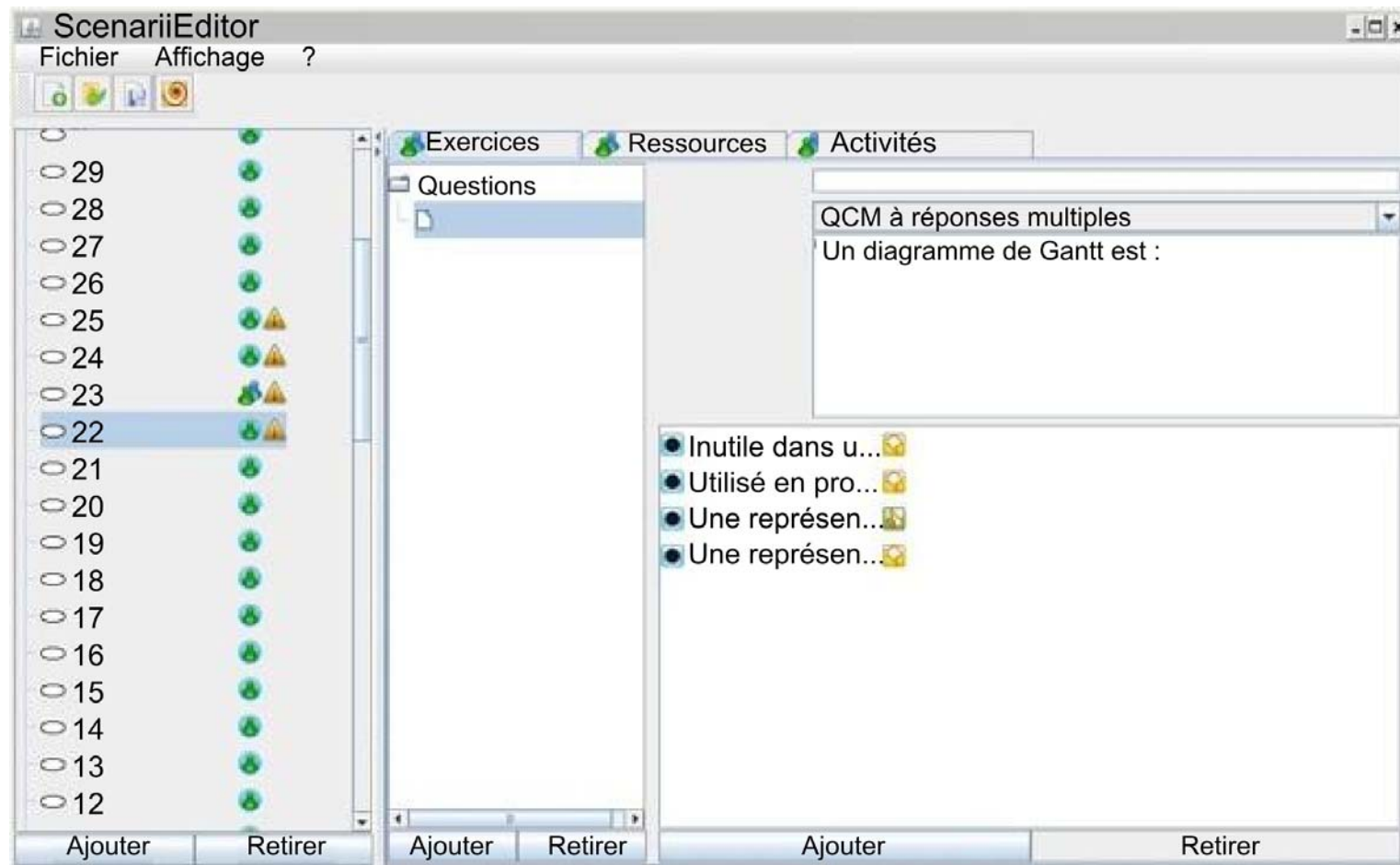
Local Groupe **Guilde** Monde Debug

Toute l'équipe Learning Adventure vous souhaite la bienvenue et espère que vous apprécierez cette plateforme pédagogique.  
Vous avez accompli un objectif.  
Vous avez accompli une quête.  
Vous obtenez ## UNSET ##.  
Vous avez accompli un objectif.  
Vous avez accompli une quête.  
Vous avez trop de Belbin.

### 3) Approche



## Learning Adventure





## Côté technique : NLTK (Python)

- Librairie open-source et gratuite
- Fonctionnalités de traitement des langues
- Interface BdD et corpus (e.g. WordNet)
- Binding avec librairies et logiciels tiers tel Prover9



1. Introduction

2. État de l'art

3. Approche

4. Conclusions et perspectives

5. Références



## 4) Conclusions et perspectives



### Approche :

Croisement entre agents conversationnels et systèmes d'évaluation de réponses libres.

## 4) Conclusions et perspectives



### Difficulté :

Absence de benchmarks, que certains expliquent par des raisons de propriété intellectuelle (Sukkarieh et Blackmore, 2009).

## 4) Conclusions et perspectives



### Enjeux :

*By year-end 2013, at least 15 percent of Fortune 1000 companies will use a virtual assistant to serve up Web self-service content to enhance their CRM offerings and service delivery.*

Gartner 2011

## 4) Conclusions et perspectives



### Éducation en ligne :

MITx, Coursera et Udacity :

- cours en ligne gratuits, qui ont déjà attiré plus de 200 000 étudiants ;
- reposent en grande partie sur des QCM pour évaluer les apprenants, à défaut de systèmes plus efficaces.

