Génie Logiciel Avancé
Cours 1 — Introduction

Stefano Zacchiroli
zack@pps.univ-paris-diderot.fr

Laboratoire PPS, Université Paris Diderot - Paris 7

URL http://upsilon.cc/zack/teaching/1112/gla/

Copyright © 2011–2012 Stefano Zacchiroli
© 2010 Yann Régis-Gianas
License Creative Commons Attribution-ShareAlike 3.0 Unported License
http://creativecommons.org/licenses/by-sa/3.0/

Sommaire

1. Qu'est-ce que le génie logiciel ?
2. Les grands principes du génie logiciel
3. Les processus de développement logiciel
4. Le cours de GLA

Sommaire

1. Qu’est-ce que le génie logiciel ?
2. Les grands principes du génie logiciel
3. Les processus de développement logiciel
4. Le cours de GLA

Qu’est-ce que le génie logiciel ?

Définition (Génie logiciel)

Le **génie logiciel** est un domaine des sciences de l’ingénieur dont l’objet d’étude est la conception, la fabrication, et la maintenance des *systèmes informatiques complexes*.

Qu’est-ce qu’un système ?

?

Qu'est-ce qu'un système ?

- Un système est un ensemble d'*éléments* interagissant entre eux suivant un certains nombres de principes et de *règles* dans le but de réaliser un *objectif*.
- La frontière d'un système est le critère d'appartenance au système.
- L'environnement est la partie du monde extérieure au système.
- Un système est souvent hiérarchisé à l'aide de *sous-systèmes*.
- Un système complexe se caractérise par :
  - sa dimension, qui nécessite la collaboration de plusieurs personnes ;
  - son évolutivité.

Exemple

une fourmilière, l'économie mondiale, le noyau Linux, ...

Qu'est-ce qu'un logiciel ?

?

Qu'est-ce qu'un logiciel ?

Définition (Logiciel)

Un logiciel est un ensemble d'entités nécessaires au fonctionnement d'un processus de traitement automatique de l'information.

Parmi ces entités, on trouve par exemple :

- des programmes (en format *code source* ou exécutables) ;
- des documentations d'utilisation ;
- des informations de configuration.

Qu'est-ce qu'un logiciel ?

Un logiciel est en général un sous-système d'un système englobant.

- Il peut interagir avec des clients, qui peuvent être :
  - des opérateurs humains (utilisateurs, administrateurs, . . .) ;
  - d'autres logiciels ;
  - des contrôleurs matériels.
- Il réalise une spécification : son comportement vérifie un ensemble de critères qui régissent ses interactions avec son environnement.

Le génie logiciel vise à garantir que :

1. la spécification répond aux besoins réels de ses clients ;
2. le logiciel respecte sa spécification ;
3. les coûts alloués pour sa réalisation sont respectés ;
4. les délais de réalisation sont respectés.

Comment spécifier un logiciel ?

Répondre à la question

Que doit faire le logiciel ?

- La spécification d'un logiciel peut prendre de nombreuses formes.
- La *complexité* et les *dimensions* de la spécification peuvent varier énormément en fonction de l'environnement d'utilisation du logiciel et des objectifs auxquels il répond.

Quelques exemples de spécifications

$$\forall I, Precondition(I) \implies \exists O, Postcondition(I, O)$$

Exemple (Un algorithme de tri)

Entrée	un tableau t
Pre	il existe une relation d'ordre sur les éléments du tableau
Sortie	un tableau u
Post	u est trié et contient exactement les mêmes éléments que t

Exemple (La partie arrière d'un compilateur)

Entrée	un arbre de syntaxe abstraite P
Pre	le programme est bien typé
Sortie	un fichier exécutable E
Post	la sémantique de E est la même que celle de P

Ce sont des spécifications (relativement) simples dont la conformité aux objectifs de leurs clients ne fait aucun doute.
(Cela ne rend pas aisée pour autant leur réalisation.)

Quelques exemples de spécifications plus complexes

Exemple (Une interface graphique)

Le modèle d'interaction avec le client est non déterministe.
Doit-on spécifier toutes les traces d'exécution possibles ?

Exemple (Un traducteur automatique)

Qu'est-ce qu'un texte anglais « bien écrit » ?

Exemple (Un logiciel « boursicoteur »)

(effectuant des achats et des ventes en bourse)
Comment établir une spécification sans y inclure un modèle du système financier ?

Exemple (Un jeu vidéo)

Comment spécifier ce qui est amusant ?

Comment fabriquer un logiciel de qualité ?

En plus du respect (essentiel) de sa spécification, la qualité d'un logiciel dépend des 4 critères suivants :

Maintenabilité Peut-on faire évoluer le logiciel ?[1]

Robustesse Le logiciel est-il sujet à des dysfonctionnements ?

Efficacité Le logiciel fait-il bon usage de ses ressources ?

Utilisabilité Est-il facile à utiliser ?

1. Un logiciel ne s'use pas. La correction d'une erreur n'est pas évolution mais un échec du concepteur.

Comment fabriquer un logiciel de qualité ?

Historiquement, il y a eu une prise de conscience dans les années 70, appelée la crise du logiciel, dû à un tournant décisif : c'est à cette époque que le coût de construction du logiciel est devenu plus important que celui de la construction du matériel.

> *The major cause of the software crisis is that the machines have become several orders of magnitude more powerful ! To put it quite bluntly : as long as there were no machines, programming was no problem at all ; when we had a few weak computers, programming became a mild problem, and now we have gigantic computers, programming has become an equally gigantic problem.*

Edsger W. Dijkstra
The Humble Programmer.
ACM Turing Lecture, 1972.
http://cs.utexas.edu/users/EWD/transcriptions/EWD03xx/EWD340.html

Comment fabriquer un logiciel de qualité ?

Pour répondre à cette crise, on a essayé d'appliquer les méthodes connues de l'ingénieur au domaine du logiciel, pour établir des méthodes fiables sur lesquelles construire une industrie du logiciel.

Il s'agit de se donner un cadre rigoureux pour :

1. Guider le développement du logiciel, de sa conception à sa livraison.
2. Contrôler les coûts, évaluer les risques et respecter les délais.
3. Établir des critères d'évaluation de la qualité d'un logiciel.

Les spécificités du logiciel

Cependant, la construction d'un logiciel diffère de celle d'un pont :

- une modification infime peut avoir des conséquences critiques ;
- les progrès technologiques très rapides peuvent rendre un logiciel caduque ;
- il est difficile de raisonner sur des programmes ;
- les domaines des entrées des logiciels sont trop grands pour le test exhaustif ;
- les défaillances des programmes sont en général dues à des erreurs humaines ;
- on ne sait pas très bien réutiliser les programmes existants ;
- chaque logiciel a son organisation et sa logique propre ;
- ...

Comment fabriquer un logiciel de qualité ?

- Le génie logiciel est un domaine en pleine évolution qui offre une grande palette d'outils et de méthodes pour parvenir à construire du logiciel de qualité.
- Aucune de ses méthodes ne s'est imposée à ce jour : il faut donc prendre du recul sur les concepts et les conseils qu'elles préconisent et utiliser son bon sens pour les adapter à chaque situation.
- Ces méthodes se distinguent principalement par :
  - leur degré de formalisme ;
  - leur champ d'application ;
  - les contraintes de qualité qu'elles ambitionnent.

Qualité du logiciel — approches formelles

Les approches formelles utilisent des outils mathématiques et des méthodes de preuve pour construire un logiciel correct par construction dont la vérification est automatisée ou assistée.

Exemple (approches formelles)

- Méthodes : méthode B, model-checking, logique de Hoare, . . .
- Outils et notations : Coq, Z, Atelier B, Why, Frama-C, . . .

- Ces méthodes sont utilisées pour développer des logiciels critiques.
- Elles correspondent au niveau le plus élevé de certification.
  - e.g. applications de la méthode B pour développer le logiciel embarqué des lignes de métro 14 (1998) et 1 à Paris

Qualité du logiciel — approches semi-formelles

Les approches semi-formelles visent à introduire un langage normalisé pour décrire le logiciel et sa spécification.

- Cependant, la *sémantique du langage de spécification n'est pas formalisée*.
- Bien que ces approches précisent le discours du concepteur si on le compare à celui décrit à l'aide du langage naturel, elles contiennent certaines ambiguïtés et n'offrent aucune garantie sur la qualité des résultats.

Exemple (approches semi-formelles)

- Méthodes : Rationale Unified Process, Merise, ...
- Outils et notations : UML, AnalyseSI, ...

Ces méthodes sont utilisées aujourd'hui par l'industrie du logiciel.

Qualité du logiciel — approches empiriques

Les approches empiriques mettent en avant un ensemble de “bonnes pratiques” qui ont fait leur preuve par l’expérience.

Exemple (approches empiriques)

- Méthodes : *unit testing*, *peer review*, relecture de code, *extreme programming*, programmation défensive, . . .
- Outils : plates-formes de test, Gerrit, gestionnaire de versions, outil de documentation automatique / *literate programming*, . . .

Sommaire

1. Qu'est-ce que le génie logiciel ?
2. Les grands principes du génie logiciel
3. Les processus de développement logiciel
4. Le cours de GLA

Les grands principes du génie logiciel

Un certain nombre de grands principes (de bon sens) se retrouvent dans toutes ces méthodes. En voici une liste proposée par Ghezzi :

1. La rigueur
2. La décomposition des problèmes en sous-problèmes indépendants
3. La modularité
4. L'abstraction
5. L'anticipation des évolutions
6. La généricité
7. La construction incrémentale

C. Ghezzi.
Fundamentals of Software Engineering.
Prentice Hall, 2nd edition, 2002.

Principe #1 — la rigueur

- Les principales sources de défaillances d'un logiciel sont d'origine humaine.
- À tout moment, il faut se questionner sur la validité de son action.
- Des outils de vérification accompagnant le développement peuvent aider à réduire les erreurs. Cette famille d'outils s'appelle CASE (*Computer Aided Software Engineering*).

Exemple

typeurs, générateurs de code, assistants de preuves, générateurs de tests, outil d'intégration continue, fuzzer, ...

Principe #2 — la décomposition en sous-problèmes

- « *Separation of concerns* » en anglais.
- Il s'agit de :
  - Décorréler les problèmes pour n'en traiter qu'un seul à la fois.
  - Simplifier les problèmes (temporairement) pour aborder leur complexité progressivement.

Principe #2 — la décomposition en sous-problèmes

Let me try to explain to you, what to my taste is characteristic for all intelligent thinking. It is, that one is willing to study in depth an aspect of one's subject matter in isolation for the sake of its own consistency, all the time knowing that one is occupying oneself only with one of the aspects. We know that a program must be correct and we can study it from that viewpoint only ; we also know that it should be efficient and we can study its efficiency on another day, so to speak. In another mood we may ask ourselves whether, and if so : why, the program is desirable. But nothing is gained—on the contrary !—by tackling these various aspects simultaneously. It is what I sometimes have called "the separation of concerns", which, even if not perfectly possible, is yet the only available technique for effective ordering of one's thoughts, that I know of. This is what I mean by "focusing one's attention upon some aspect" : it does not mean ignoring the other aspects, it is just doing justice to the fact that from this aspect's point of view, the other is irrelevant. It is being one- and multiple-track minded simultaneously.

Edsger W. Dijkstra
On the role of scientific thought.
http://cs.utexas.edu/users/EWD/transcriptions/EWD04xx/EWD447.html

Principe #2 — la décomposition en sous-problèmes

Exemple (Comment acheminer un email de façon sûr à travers un réseau ?)

Décomposition en couches utilisée sur Internet :

SMTP
: protocole de la couche application (sûr, grâce au *store and forward*) qui suppose une couche de transport de paquet sûr.

TCP
: protocole de la couche transport permettant de s'assurer que tous les paquets arrivent, même si le réseau peut perdre des paquets.

Principe #2 — la décomposition en sous-problèmes

Exemple (Comment créer dynamiquement une page internet pour visualiser et modifier le contenu d'une base donnée sans la corrompre ?)

Décomposition en trois composants :

- **Modèle** son rôle est gérer le stockage des données.
- **Vue** son rôle est formatter les données.
- **Contrôleur** son rôle est de n'autoriser que les modifications correctes.

Principe #3 — la modularité

- C'est une instance cruciale du principe de décomposition des problèmes.
- Il s'agit de partitionner le logiciel en modules qui :
  - ont une cohérence interne (des invariants) ;
  - possèdent une interface ne divulgant sur le contenu du module que ce qui est strictement nécessaire aux modules clients.
- L'évolution de l'interface est indépendante de celle de l'implémentation du module.
- Les choix d'implémentation sont indépendants de l'utilisation du module.
- Ce mécanisme s'appelle le camouflage de l'information (*information hiding*).

D. L. Parnas.
On the criteria to be used in decomposing systems into modules.
Communications of the ACM. Vol. 15 Issue 12, 1972

Principe #4 — l'abstraction

- C'est encore une instance du principe de décomposition des problèmes.
- Il s'agit d'exhiber des concepts généraux regroupant un certain nombre de cas particuliers et de raisonner sur ces concepts généraux plutôt que sur chacun des cas particuliers.
- Le fait de fixer la bonne granularité de détails permet :
  - de raisonner plus efficacement ;
  - de factoriser le travail en instanciant le raisonnement général sur chaque cas particulier.

Exemple (Support dans les langages de programmation)

les classes abstraites dans les langages à objets, le polymorphisme de Caml et le *generics* de Java, les fonctions d'ordre supérieur, les foncteurs de Caml et le *templates* de C++, . . .

Principe #5 — l'anticipation des évolutions

- Un logiciel a un cycle de vie plus complexe que l'habituel cycle « commande → spécification → production → livraison ».
- La maintenance est la gestion des évolutions du logiciel.
- Il est primordial de prévoir les évolutions possibles d'un logiciel pour que la maintenance soit la plus efficace possible.
- Pour cela, il faut s'assurer que les modifications à effectuer soient le plus locales possibles.
- Ces modifications ne devraient pas être *intrusives* car les modifications du produit existant remettent en cause ses précédentes validations.
- Concevoir un système suffisamment riche pour que l'on puisse le modifier **incrémentalement** est l'idéal.

Principe #6 — la généricité

Figure: *template*

- Un logiciel réutilisable a beaucoup plus de valeur qu'un composant dédié.
- Un composant est générique lorsqu'il est adaptable.

Principe #7 — la construction incrémentale

- Un développement logiciel a plus de chances d'aboutir si il suit une cheminement incrémental (*baby-steps*).

Exemple

Laquelle de ses deux méthodes de programmation est la plus efficace ?

1. Écrire l'ensemble du code source d'un programme et compiler.
2. Écrire le code source d'une fonction ou module, le compiler, et passer à la suivante.

Pourquoi ?

À propos des principes

- Vous devez avoir en tête ces principes : ils se retrouvent dans toutes les méthodes et outils que nous allons aborder.

1. La rigueur
2. La décomposition des problèmes en sous-problèmes indépendants
3. La modularité
4. L'abstraction
5. L'anticipation des évolutions
6. La généricité
7. La construction incrémentale

Sommaire

1 Qu'est-ce que le génie logiciel ?

2 Les grands principes du génie logiciel

3 Les processus de développement logiciel

4 Le cours de GLA

Qu'est-ce qu'un processus ?

Définition (processus de développement logiciel)

Un processus de développement logiciel est un ensemble (structuré) d'activités que conduisent à la production d'un logiciel

- Il n'existe pas de processus idéal.
- La plupart des entreprises adapte les processus existants à leurs besoins.
- Ces besoins varient en fonction du domaine, des contraintes de qualité, des personnes impliquées.

- Ce qui est essentiel, c'est de comprendre quel est son rôle dans ce processus et d'en saisir les rouages.
- L'étude et la pratique de processus existants doit vous permettre de vous forger un regard affûté (et même critique) sur ces processus.

Activités du développement logiciel

Les activités des processus de développement logiciels se regroupent en 5 grandes catégories :

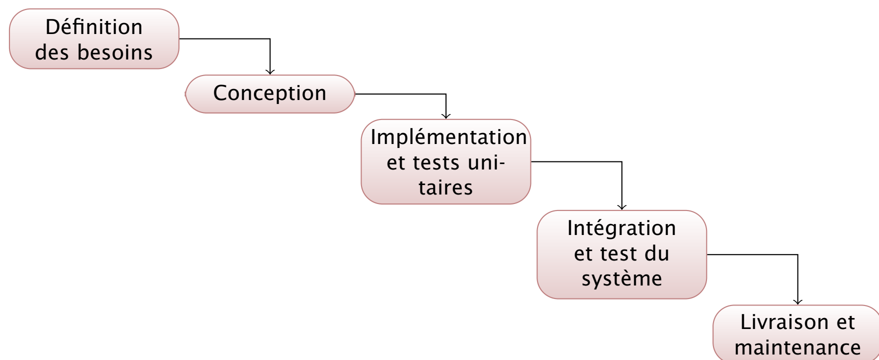
1. La spécification du logiciel définit ses fonctionnalités et leurs contraintes.
2. La conception . . .
3. . . . et l'implémentation sont chargées de réaliser le logiciel, en conformité avec sa spécification.
4. La validation s'assure effectivement du respect de la spécification par le logiciel produit.
5. L'évolution adapte le logiciel aux besoins futurs de ses clients.

Schéma général d'un processus de développement

- Il est très rare d'appliquer un processus comme une unique séquence des 5 activités précédentes.
- En effet, ce serait à l'encontre du principe d'incrémentalité.
- En général, un logiciel complet est le fruit de plusieurs itérations.
- Chaque itération contient les 5 activités de spécification, conception, implémentation, validation et évolution.

Il existe différents modèles de processus qui organisent de façon différentes ces activités, entre eux : le modèle en *cascade*, le modèle de *développement évolutif* et le modèle de *développement par composants*.

Modèle en cascade

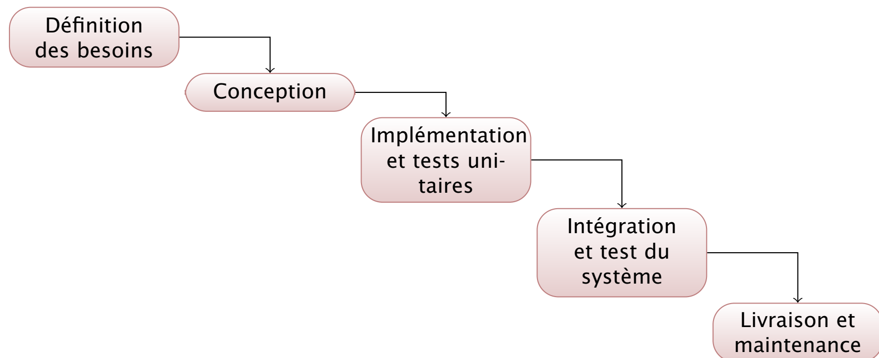


- Chaque phase doit se terminer pour commencer la suivante.
- Des documents sont produits pour concrétiser la réalisation de chaque phase.

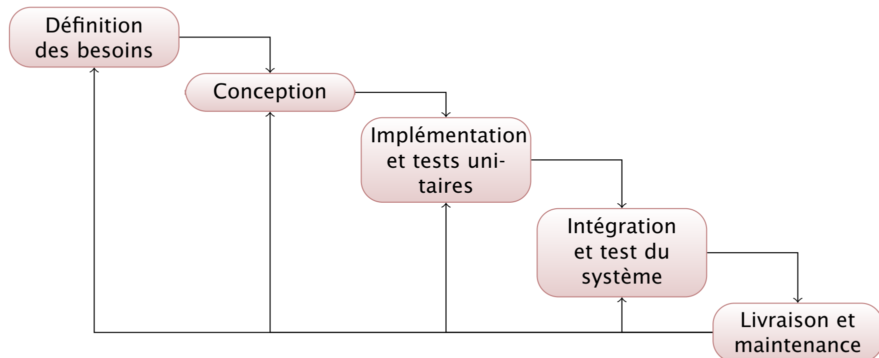
Modèle en cascade — caractéristiques

- Le modèle en cascade est hérité des méthodes classiques d'ingénierie.
  - Il s'adapte donc bien dans un contexte où le logiciel fait partie d'un système complexe englobant.
- La production de documents entre chaque phase améliore le suivi du projet.
- Lorsqu'une erreur a été commise dans une phase et qu'elle est détectée dans une phase suivante, il faut faire remonter cette information dans la phase incriminée et recommencer le processus à partir de celle-ci. On doit alors reproduire de nouveaux documents . . .
- Ce modèle de processus impose donc une importante réflexion sur les choix faits en amont car le coût de la correction d'une erreur est important.
  - Typique d'un développement industriel pour lequel les coûts de la construction du produit sont trop importants pour se permettre une erreur de choix de conception.

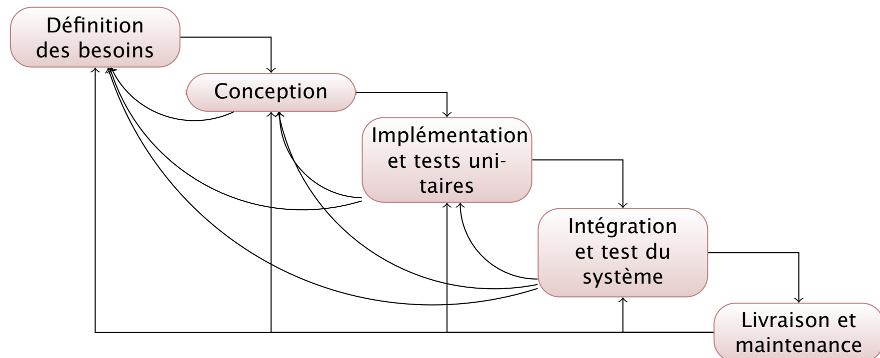
Modèle en cascade — “*feedback loop*”



Modèle en cascade — “*feedback loop*”



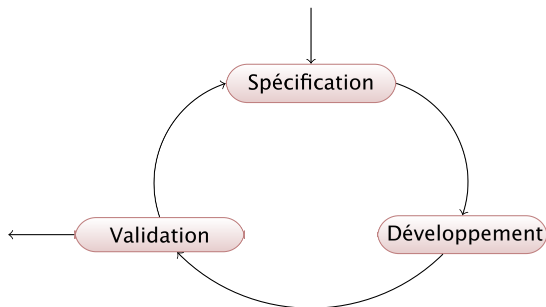
Modèle en cascade — “*feedback loop*”



Modèle en cascade — critique

- Le modèle en cascade rend coûteux le développement itératif puisque la rédaction des documents de validation de chaque phase demande beaucoup de travail.
- Ce modèle est inadapté au développement de systèmes dont la spécification est difficile à formuler *a priori*.

Modèle de développement évolutif



- Ces trois activités sont entrelacées.
- Un prototype est écrit rapidement et est confronté à l'utilisateur.
- En fonction du résultat, on raffine la spécification.
- On reprend le prototype ou on le réécrit jusqu'à l'obtention du système final.

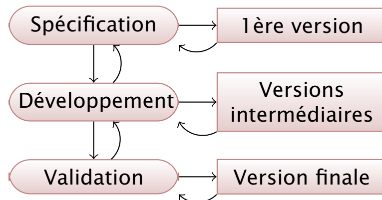
Modèle de développement évolutif — caractéristiques

- Ce modèle augmente les chances de répondre aux besoins de l'utilisateur car il permet de les comprendre plus rapidement.
  - *Are we building the right product ?*
- Il remplit le critère d'incrémentalité.
- Ce modèle ne dispense d'écrire la spécification du système car il faut s'assurer que l'implémentation est correct.
  - *Are we building the product right ?*
- C'est un processus particulièrement adapté aux projets de taille moyenne (inférieur à 100 000 lignes de code) comme par exemple des grosses applications Web ou encore les solutions intégrées pour les petites entreprises.

Modèle de développement évolutif — critique

- Il est plus difficile de gérer un projet utilisant ce modèle car la visibilité de l'avancement du développement est peu clair.
  - Dans ce cadre, encore plus que dans un autre, un chef de projet doit aussi être un bon programmeur puisqu'il doit être capable de se faire une idée de l'état du système en observant le développement (possiblement chaotique) des prototypes.
- Il est plus difficile de structurer correctement le logiciel (définir de bonnes abstractions, modulariser efficacement) car les prototypes sont par définition des produits "bricolés".
- Le coût en termes de tests et de validation du produit final peuvent être très importants.
  - Des approches mixtes intégrant modèle de développement évolutif pour produire un premier prototype validé et un modèle en cascade pour reconstruire correctement un produit final constituent en général de bons compromis.

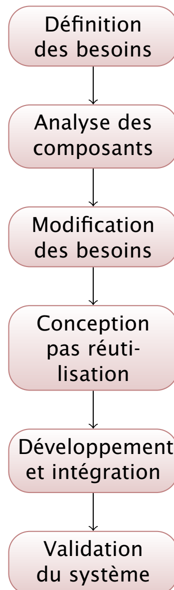
Modèle de développement à livraison incrémentale



Une approche à mi-chemin entre les modèles cascade et évolutif s'appuie sur une livraison incrémentale du produit.

- On hiérarchise les besoins du client en termes de priorité.
- Chaque itération du modèle vise à obtenir un ensemble de fonctionnalités par ordre de priorité.
- Traiter les parties les plus critiques du système en premier minimise les risques d'inadéquation avec le produit final.
- Cependant, il se peut que les choix pris en amont, trop focalisés sur ce noyau de fonctionnalités, compromettent le développement des fonctionnalités secondaires.

Modèle de développement par composants



Modèle de dével. par composants — caractéristiques

- Ce modèle vise à développer un logiciel en grande partie à l'aide d'une base de composants génériques pré-existants.
- L'élaboration de la spécification est dirigée par cette base : une fonctionnalité est proposée à l'utilisateur en fonction de sa facilité à l'obtenir à l'aide d'un composant existant.
  - Situation typique chez les sociétés de services (hébergement de serveurs, déploiement automatique de site Web, . . .).
- Ce modèle permet d'obtenir rapidement des produits de bonne qualité puisqu'ils sont construits à partir de composants qui ont fait leur preuve.
- Le travail d'intégration peut s'appuyer sur des outils dirigés par des descriptions de haut niveau du système qui génèrent le code de "glue" par exemple.

Modèle de dével. par composants — critique

- Le principal défaut de ce modèle est de ne pas construire un produit adapté aux besoins du client.
  - Un travail complexe de configuration et d'adaptation peut être nécessaire.

Sommaire

1 Qu'est-ce que le génie logiciel ?

2 Les grands principes du génie logiciel

3 Les processus de développement logiciel

4 Le cours de GLA

Objectifs du cours

- Ce cours a pour but de vous familiariser avec les futures structures de votre vie professionnelle et de vous donner les outils de vous adapter à la situation, nécessairement singulière, dans laquelle vous serez acteurs.
- Il a aussi pour objectif de développer vos capacités d'analyse de problèmes de conception logiciel.

Contenu du cours

1. Le génie logiciel et ses grands principes
2. Les principaux processus de développement logiciel
   - Cascade, évolutif, incrémentale, etc.
3. Spécification logiciel
   - Panoramiques des méthodes des spécification
   - Faisabilité et analyse des besoins
4. Introduction à la spécification formelle du logiciel
   - Méthodes formelles pour la spécification logiciel, des exemples
5. La modélisation à objet et le langage UML
   - Vues des cas d'utilisation, d'architecture, statiques et dynamiques
6. Patrons de conception à objet
   - Patrons de création, structure, et comportement
   - Les anti-patrons de conception
7. Méthodes agiles et extreme programming
8. Maîtriser le changement
   - Introduction aux techniques de testing et de contrôle de version

Calendrier

- cours : mercredi 10h30-12h30, amphi 8C
- TD (groupe A) : mardi 11h30-13h30, salle 473F
  - en salle machine 548C le 28/02, 13/03, 27/03, 03/04
- TD (groupe B) : mercredi 12h30-14h30, salle 478F
  - en salle machine 557C le 29/02, 14/03, 28/03, 04/04, 02/05

1e TD : 14/15 février 2012

- cours : Stefano Zacchiroli et Delia Kesner
- TD : Stefano Zacchiroli, Delia Kesner, Fabien Renaud

Validation

- Le cours est validé par un projet et par un examen (50/50).
  - le projet n'est pas du contrôle continu, il est donc obligatoire
- Le projet consiste à développer un logiciel de façon collaboratif en équipe, en utilisant les méthodes et outils de génie logiciel que nous découvrirons
  - voir les TDs

Ressources

Page web du cours

`http://upsilon.cc/~zack/teaching/1112/gla/`

L'inscription aux listes de diffusion est obligatoire :

- `m1gl` — annonces
- `m1gl-projet` — discussions sur le projet

voir la page du cours pour plus des liens et d'instructions.

Bibliographie

Ian Sommerville
Software Engineering.
Pearson, 9th edition, 2010.

Erich Gamma, Richard Helm, Ralph Johnson, John M. Vlissides
Design Patterns : Elements of Reusable Object-Oriented Software.
Addison-Wesley, 1994.

Kent Beck
Extreme Programming Explained : Embrace Change.
Addison-Wesley, 2nd edition, 2004.