

LINQ To Objects

Morre Thierry

Cognitic

thierry.morre@cognitic.be

Sommaire

- Introduction
- LINQ => Query ou pas!
- IEnumerable<T> et séquence
- Opérateurs de requête standard
- Modes d'exécutions
- Support de l'expression de requête
- Les opérateurs courants
- Utilisation de clés composites
- Multiples clauses « from »

Introduction

LINQ To Objects fait référence à l'utilisation directe de requêtes LINQ avec n'importe quelle collection « IEnumerable » et « IEnumerable<T> » telles que « List<T> », « Array » ou « Dictionary<Tkey, Tvalue> » définies par l'utilisateur ou retournées par une API du Framework .NET.

Auparavant, nous devions écrire des boucles complexes pour spécifier comment récupérer des données d'une collection. Maintenant grâce à LINQ, nous pourrons écrire du code déclaratif qui décrira exactement ce que nous voudrions.

Ces requêtes offrent trois principaux avantages par rapport aux boucles :

- Elles sont plus concises et lisibles
- Elles fournissent des fonctions puissantes de filtrage, de classement et de regroupement avec un minimum de code.
- Elles peuvent être appliquées à d'autres sources de données avec peu ou pas de changement.

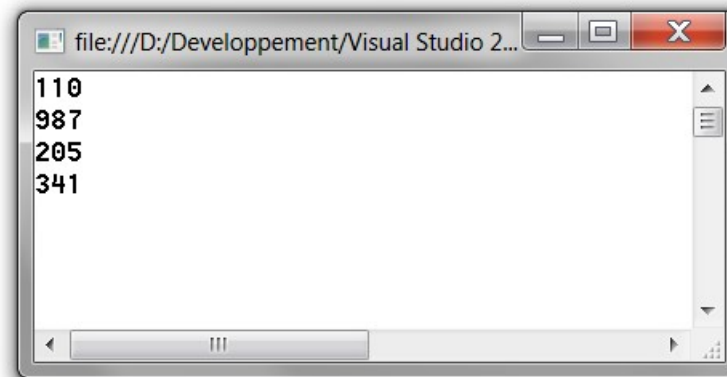
En gros, plus l'opération sur les données sera complexe, plus nous verrons les avantages d'utiliser LINQ.

```
IEnumerable<Contact> QueryResult = from Contact c in Contacts
                                   where c.Email.Substring(0, c.Email.IndexOf('@')).Length == (c.Nom.Length + c.Prenom.Length + 1)
                                   orderby c.Prenom, c.Nom
                                   select c;
```

LINQ => Query ou pas!

LINQ étant l'abréviation de « **Language Integrated Query** », nous pourrions penser qu'il se limite à l'interrogation de données. Mais en réalité, son utilisation peut aller bien au-delà.

```
string[] ValuesAsStrings = { "110", "987", "205", "341" };  
int[] ints = ValuesAsStrings.Select(s => Int32.Parse(s)).ToArray();  
  
foreach (int i in ints)  
{  
    Console.WriteLine(i);  
}  
  
Console.ReadLine();
```



La conversion de ce tableau de « string » en « int » pourrait être plus simple?

IEnumerable<T> et séquence

Cette interface a été implémentée par les tableaux et les classes de collections génériques de C# 2.0. Elle permet, de par son fonctionnement, d'énumérer les éléments d'une collection.

Une séquence est un terme logique d'une collection implémentant l'interface « IEnumerable<T> ». En d'autres termes, si nous avons une variable de type « IEnumerable<string> », nous pourrions dire que nous avons une séquence de « string ».

Nous allons voir, par la suite que la plupart des « Opérateurs de requête standard » sont prototypées de façon à retourner une séquence (IEnumerable<T>).

Contacts.Select(

▲ 1 of 2 ▼ (extension) IEnumerable<TResult> IEnumerable<Contact>.Select(Func<Contact, TResult> selector)

Projects each element of a sequence into a new form.

selector: A transform function to apply to each element.

Opérateurs de requête standard

Les « opérateurs de requête standard » sont les méthodes qui composent le modèle LINQ.

La plupart de ses méthodes fonctionnent sur des séquences et fournissent des fonctions de requête, notamment le filtrage (Where), la projection (Select), l'agrégation (Sum), le tri (OrderBy), etc..

Elles sont définies, dans la classe « Enumerable », comme méthodes d'extension à l'interface « IEnumerable », cela implique qu'elle peuvent être appelées sur n'importe quelle collection générique comme méthode d'instance.

Ces opérateurs sont regroupés par fonctionnalité :

- Agrégation
- Concaténation
- Conversion
- Egalité
- Élément
- Ensemble
- Filtrage
- Génération
- Jointure
- Partitionnement
- Projection
- Quantificateur
- Regroupement
- Tri

En annexe vous trouverez la liste complète de ces opérateurs (.NET 3.5) par ordre alphabétique.

Modes d'exécutions

Les opérateurs de requêtes standard se divise en 2 modes d'exécution.

Immédiat :

L'exécution immédiate signifie que la source de données est lue et que l'opération est effectuée au point où la requête est déclarée dans le code. Ce mode d'exécution vise tous les opérateurs de requête standard qui retournent un résultat unique et non énumérable comme les opérateurs d'agrégation, d'élément, etc.

Différé :

L'exécution différée signifie que l'opération n'est pas effectuée au point où la requête est déclarée dans le code mais qu'elle le sera uniquement que lorsque la variable de requête est énumérée (boucle foreach par exemple).

Cela signifie que les résultats de l'exécution de la requête dépendent du contenu de la source de données lorsque la requête est exécutée plutôt que lorsqu'elle est définie. Si la variable de requête est énumérée plusieurs fois, les résultats peuvent s'avérer différents chaque fois. Presque tous les opérateurs de requête standard dont le type de retour est `IEnumerable<T>` s'exécutent de manière différée.

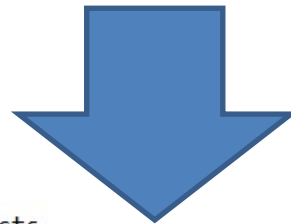
Support de l'expression de requête

Certains des opérateurs de requêtes standard, les plus courants, possèdent une syntaxe de mots clé du langage C# qui leur permet d'être appelés dans le cadre d'une expression de requête.

Par défaut, pour utiliser LINQ, nous devrions appeler ces opérateurs qui sont des méthodes. L'expression de requête nous offre une forme plus lisible que son équivalent fondé sur des méthodes.

Les clauses de requêtes sont traduites en appels de méthodes lors de la compilations.

```
IEnumerable<Contact> QueryResult = from Contact c in Contacts
                                   where c.Email.Substring(0, c.Email.IndexOf('@')).Length == (c.Nom.Length + c.Prenom.Length + 1)
                                   orderby c.Prenom, c.Nom
                                   select c;
```



```
IEnumerable<Contact> QueryResult = Contacts
    .Where(c => c.Email.Substring(0, c.Email.IndexOf('@')).Length == (c.Nom.Length + c.Prenom.Length + 1))
    .OrderBy(c => c.Prenom)
    .ThenBy(c => c.Nom);
```

Une expression de requête doit commencer par une clause « from » et doit se terminer par une clause « select » ou « group ». Entre la première clause from et la dernière clause select ou group, elle peut contenir une ou plusieurs clauses facultatives : where, orderby, join, let et même d'autres clauses from supplémentaires.

Nous pourrions également utiliser le mot clé « into » pour que le résultat d'une clause « join » ou « group » puisse servir de source pour des clauses de requête supplémentaires dans la même expression de requête.

Support de l'expression de requête

Liste des Opérateurs supportant l'expression de requête en C#

Opérateur	Syntaxe d'expression de requête
Cast	'from int n in numbers' explicitement typé
GroupBy	'group' ... by ou 'group ... by ... into ...'
GroupJoin	'join ... in ... on ... equals ... into ...'
Join	'join ... in ... on ... equals ...'
OrderBy	'order by ...'
OrderByDescending	'orderby ... descending'
Select	'select'
SelectMany	Plusieurs clauses 'from'
ThenBy	'orderby ..., ...'
ThenByDescending	'orderby ..., ... descending'
Where	'where'

Les opérateurs courants

Nous allons à présent faire le tour des opérateurs les plus souvent utilisés en LINQ. Les exemples qui vont suivre seront exécutés sur une liste générique de « Contact ».

```
public class Contact
{
    public string Nom { get; set; }
    public string Prenom { get; set; }
    public string Email { get; set; }
    public int AnneeDeNaissance { get; set; }
}
```

```
List<Contact> Contacts = new List<Contact>();
Contacts.AddRange(new Contact[] {
    new Contact(){ Nom = "Person", Prenom="Michael", Email="michael.person@cognitic.be", AnneeDeNaissance = 1982 },
    new Contact(){ Nom = "Morre", Prenom="Thierry", Email="thierry.morre@cognitic.be", AnneeDeNaissance = 1974 },
    new Contact(){ Nom = "Dupuis", Prenom="Thierry", Email="thierry.dupuis@cognitic.be", AnneeDeNaissance = 1988 },
    new Contact(){ Nom = "Faulkner", Prenom="Stéphane", Email="stephane.faulkner@cognitic.be", AnneeDeNaissance = 1969 },
    new Contact(){ Nom = "Selleck", Prenom = "Tom", Email = "tom.selleck@imdb.com", AnneeDeNaissance = 1945 },
    new Contact(){ Nom = "Anderson", Prenom = "Richard Dean", Email = "richard.dean.anderson@imdb.com", AnneeDeNaissance = 1950 },
    new Contact(){ Nom = "Bullock", Prenom = "Sandra", Email = "sandra.bullock@imdb.com", AnneeDeNaissance = 1964 },
    new Contact(){ Nom = "Peppard", Prenom = "George", Email = "peppard.george@ateam.com", AnneeDeNaissance = 1928 },
    new Contact(){ Nom = "Estevez", Prenom = "Emilio", Email = "emilio.estevez@breakfirstclub.com", AnneeDeNaissance = 1962 },
    new Contact(){ Nom = "Moore", Prenom = "Demi", Email = "demi.moore@imdb.com", AnneeDeNaissance = 1962 },
    new Contact(){ Nom = "Willis", Prenom = "Bruce", Email = "bruce.willis@diehard.com", AnneeDeNaissance = 1955 },
});
```

Les opérateurs courants

Voici les opérateurs que nous allons voir :

- Opérateurs « Cast<T> » & « OfType<T> »
- Opérateur « Where »
- Opérateur « Select »
- Opérateur « Distinct »
- Opérateur « SingleOrDefault »
- Opérateur « FirstOrDefault »
- Opérateurs « OrderBy[Descending] »
- Opérateurs « ThenBy[Descending] »
- Opérateur « Count » & « LongCount »
- Opérateurs « Min » & « Max »
- Opérateurs « Sum » & « Average »
- Opérateur « GroupBy »
- Opérateur « Join »
- Opérateur « GroupJoin »

Dans certains cas, les opérateurs ont des surcharges. Pour ne pas rallonger inutilement ce module, je me contenterai des prototypes de base.

Vous trouverez l'ensemble des surcharges sur le site « MSDN » ou dans la classe « Enumerable ».

J'ajouterai la requête sous forme d'« expression de requête », lorsque l'opérateur supportera cette syntaxe.

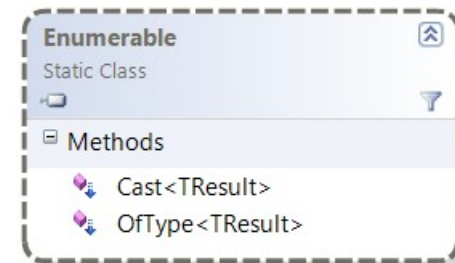
Opérateurs « Cast<T> » & « OfType<T> »

La grande majorité des opérateurs de requête LINQ ne peut être utilisée que sur des collections qui implémentent l'interface « IEnumerable<T> ». Or, aucune des collections présente dans l'espace de nom « System.Collections », telles que Array, ArrayList ou Hashtable, n'implémentent cette interface.

Mais alors comment utiliser LINQ avec ces collections?

En regardant de plus près la classe « Enumerable » qui implémente les méthodes d'extension de LINQ. Nous retrouvons 2 méthodes qui ne sont pas des méthodes d'extension sur l'interface « IEnumerable<T> » mais sur « IEnumerable ».

Ces 2 opérateurs sont là pour transformer une collection « IEnumerable » en une séquence « IEnumerable<T> »



Opérateurs « Cast<T> » & « OfType<T> »

Cependant, ces opérateurs sont quelque peu différents l'un de l'autre.

Cast<T>: `public static IEnumerable<TResult> Cast<TResult>(this IEnumerable source);`

Cet opérateur va tenter de convertir tous les éléments de la collection en une séquence de type 'T'. Si celle-ci n'y parvient pas, elle lèvera une exception.

```
//Transforme List<Contact> (IEnumerable<T>) en ArrayList (IEnumerable)
ArrayList MyArrayList = new ArrayList(Contacts.ToArray());

IEnumerable<Contact> Result = MyArrayList.Cast<Contact>();
```

Cet opérateur supporte la syntaxe d'expression de requête. En spécifiant explicitement le type à utiliser.

```
//Transforme List<Contact> (IEnumerable<T>) en ArrayList (IEnumerable)
ArrayList MyArrayList = new ArrayList(Contacts.ToArray());

IEnumerable<Contact> Result = from Contact c in MyArrayList
                             select c;
```

Opérateurs « Cast<T> » & « OfType<T> »

Celles-ci sont quelque peu différentes l'une de l'autre.

OfType<T>: `public static IEnumerable<TResult> OfType<TResult>(this IEnumerable source);`

Cet opérateur, quant à lui, va filtrer les valeurs en fonction de leur capacité à être castées dans le type spécifié.

```
//Transforme List<Contact> (IEnumerable<T>) en ArrayList (IEnumerable)
ArrayList MyArrayList = new ArrayList(Contacts.ToArray());

IEnumerable<Contact> Result = MyArrayList.OfType<Contact>();
```

Cet opérateur n'étant pas supporté par l'expression de requête, voici comment spécifier que l'on utilise « OfType<T> » plutôt que « Cast<T> ».

```
//Transforme List<Contact> (IEnumerable<T>) en ArrayList (IEnumerable)
ArrayList MyArrayList = new ArrayList(Contacts.ToArray());

//Attention 2 requêtes LINQ imbriquées
IEnumerable<Contact> Result = from c in MyArrayList.OfType<Contact>()
                             select c;
```

Opérateurs « Cast<T> » & « OfType<T> »

Préférons « OfType<T> » à « Cast<T> » :

Bien que ces deux opérateurs soient utilisable sur une collection héritées (System.Collection), « Cast<T> » nécessite que tous les éléments aient le type attendu.

Pour éviter de générer des exceptions en cas d'incompatibilité de types, préférons-lui l'opérateur « OfType<T> ». Par son intermédiaire, seuls les objets pouvant être casté dans le type attendu seront stockés dans la séquence.

Les autres seront purement et simplement ignoré.

Opérateur « Where »

```
public static IEnumerable<TSource> Where<TSource>(this IEnumerable<TSource> source, Func<TSource, bool> predicate);
```

L'opérateur « Where » est utilisé pour filtrer une séquence.

Cet opérateur utilise le « delegate » générique « Func<T, bool> » vu dans le chapitre précédent.

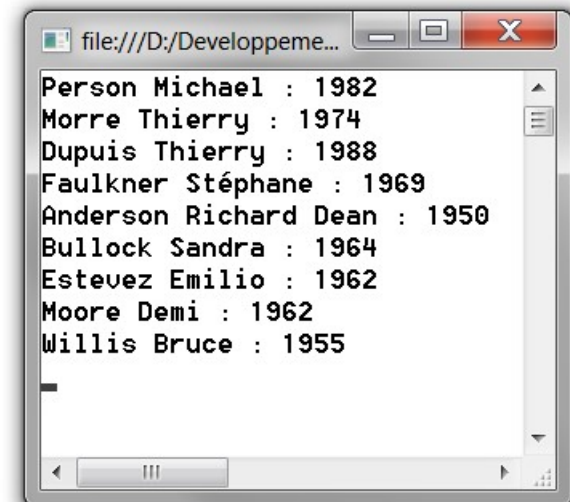
```
IEnumerable<Contact> QueryResult = Contacts.Where(c => c.AnneeDeNaissance >= 1950);
```

```
foreach (Contact c in QueryResult)
{
    Console.WriteLine("{0} {1} : {2}", c.Nom, c.Prenom, c.AnneeDeNaissance);
}
```

Expression de requête :

```
IEnumerable<Contact> QueryResult = from c in Contacts
                                   where c.AnneeDeNaissance >= 1950
                                   select c;

foreach (Contact c in QueryResult)
{
    Console.WriteLine("{0} {1} : {2}", c.Nom, c.Prenom, c.AnneeDeNaissance);
}
```



Opérateur « Select »

```
public static IEnumerable<TR> Select<T, TR>(this IEnumerable<T> source, Func<T, TR> selector);
```

L'opérateur « Select » est utilisé pour retourner une séquence d'éléments sélectionnés dans la séquence d'entrée ou à partir d'une portions de la séquence d'entrée.

Dans les requêtes LINQ basées sur les appels de méthodes, l'opérateur « Select » est facultatif et sert principalement à sélectionner une portion, en termes de Propriétés, des objets en entrée et de créer de nouveaux types d'objets (types anonymes).

Dans les expressions de requête, que l'on crée un type anonyme ou non, celui-ci est obligatoire.

```
//Sélectionne tous les contacts et retourne une séquence de type "Contact"  
IEnumerable<Contact> QueryResult = from c in Contacts  
    select c;
```

Opérateur « Select »

Création d'un type anonyme :

Pour créer un type anonyme, nous devons joindre à notre select le mot clé « new ».

```
var QueryResult = Contacts.Select(c => new { Nom = c.Nom, Courriel = c.Email });

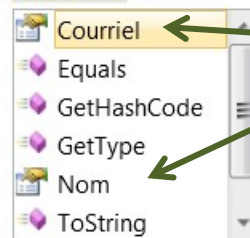
foreach (var c in QueryResult)
{
    Console.WriteLine("{0} : {1}", c.Nom, c.Courriel);
}
```

Expression de requête :

```
var QueryResult = from c in Contacts
                  select new { Nom = c.Nom, Courriel = c.Email };

foreach (var c in QueryResult)
{
    Console.WriteLine("{0} : {1}", c.Nom, c.Courriel);
}

Console.ReadLine();
```



Création à la volée de propriété

Opérateur « Distinct »

```
public static IEnumerable<T> Distinct<T>(this IEnumerable<T> source);
```

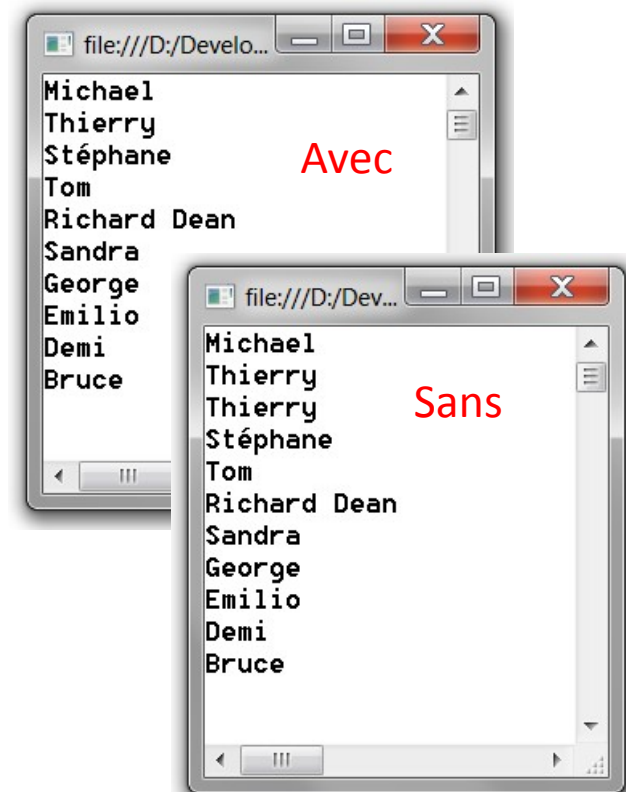
L'opérateur « Distinct » permet de supprimer les doublons dans une séquence.

```
var QueryResult = Contacts.Select(c => new { Prenom = c.Prenom }).Distinct();  
  
foreach (var c in QueryResult)  
{  
    Console.WriteLine("{0}", c.Prenom);  
}
```

Utilisation avec une expression de requête :

```
var QueryResult = (from c in Contacts  
                   select new { Prenom = c.Prenom }).Distinct();  
  
foreach (var c in QueryResult)  
{  
    Console.WriteLine("{0}", c.Prenom);  
}
```

Le distinct se fait sur la valeur de l'ensemble des propriétés.



Opérateur « SingleOrDefault »

```
public static TSource SingleOrDefault<TSource>(this IEnumerable<TSource> source);
```

L'opérateur « SingleOrDefault » retourne un élément unique d'une séquence de type 'T', si la valeur est introuvable, il retourne la valeur de « default(T) ».

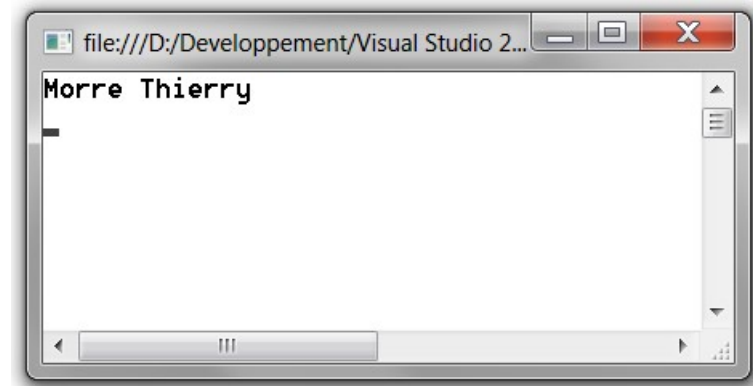
```
Contact QueryResult = Contacts
    .Where(c => c.Prenom.Equals("Thierry") && c.Nom.Equals("Morre"))
    .SingleOrDefault();

Console.WriteLine("{0} {1}", QueryResult.Nom, QueryResult.Prenom);
```

Utilisation avec une expression de requête :

```
Contact QueryResult = (from c in Contacts
    where c.Prenom.Equals("Thierry") && c.Nom.Equals("Morre")
    select c).SingleOrDefault();

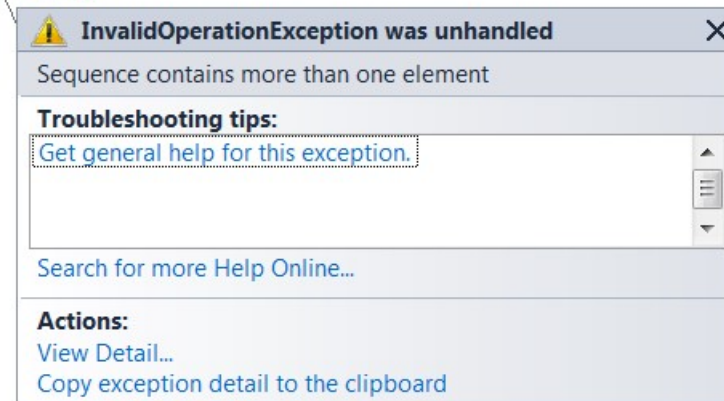
Console.WriteLine("{0} {1}", QueryResult.Nom, QueryResult.Prenom);
```



Opérateur « SingleOrDefault »

Cependant, si la requête retourne plus d'un éléments, l'opérateur « SingleOrDefault » lèvera une exception.

```
Contact QueryResult = (from c in Contacts  
                        where c.Prenom.Equals("Thierry")  
                        select c).SingleOrDefault();
```



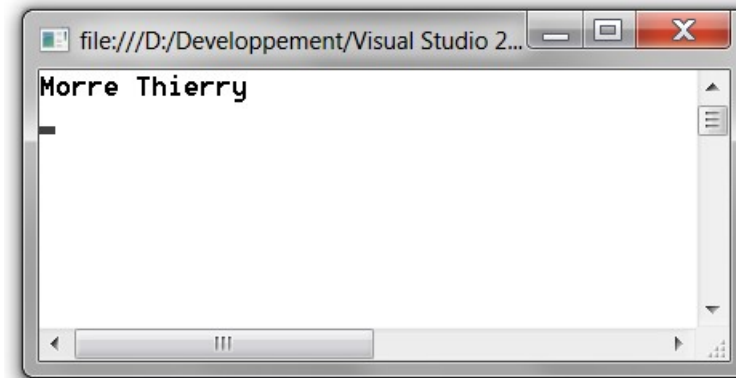
Dans ce cas de figure et afin d'éviter l'erreur, il est conseillé d'utiliser « FirstOrDefault ».

Opérateur « FirstOrDefault »

```
public static TSource FirstOrDefault<TSource>(this IEnumerable<TSource> source);
```

L'opérateur « FirstOrDefault » retourne le premier élément d'une séquence de type 'T' ou la valeur de « default(T) » dans le cas où aucune valeur ne serait trouvée.

```
Contact QueryResult = Contacts  
    .Where(c => c.Prenom.Equals("Thierry"))  
    .FirstOrDefault();  
  
if(QueryResult != null)  
    Console.WriteLine("{0} {1}", QueryResult.Nom, QueryResult.Prenom);
```



Opérateurs « OrderBy[Descending] »

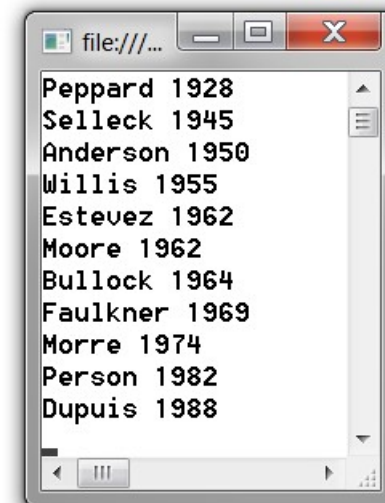
```
public static IOrderedEnumerable<T> OrderBy<T, TKey>(this IEnumerable<T> source, Func<T, TKey> keySelector);  
public static IOrderedEnumerable<T> OrderByDescending<T, TKey>(this IEnumerable<T> source, Func<T, TKey> keySelector);
```

L'opérateur « OrderBy » trie par ordre croissant les éléments d'une séquence sur base d'une clé (propriété). Pour un ordre décroissant, nous devrions utiliser « OrderByDescending ». Si nous souhaitons trier sur base de plusieurs clés, ces opérateurs doivent être les premiers à être utilisés.

```
IEnumerable<Contact> QueryResult = Contacts.OrderBy(c => c.AnneeDeNaissance);  
  
foreach (Contact c in QueryResult)  
{  
    Console.WriteLine("{0} {1}", c.Nom, c.AnneeDeNaissance);  
}
```

Expression de requête :

```
IEnumerable<Contact> QueryResult = from c in Contacts  
                                   orderby c.AnneeDeNaissance  
                                   select c;  
  
foreach (Contact c in QueryResult)  
{  
    Console.WriteLine("{0} {1}", c.Nom, c.AnneeDeNaissance);  
}
```



Opérateurs « ThenBy[Descending] »

```
public static IOrderedEnumerable<T> ThenBy<T, TKey>(this IOrderedEnumerable<T> source, Func<T, TKey> keySelector);  
public static IOrderedEnumerable<T> ThenByDescending<T, TKey>(this IOrderedEnumerable<T> source, Func<T, TKey> keySelector);
```

Les opérateurs « ThenBy » (Croissant) et « ThenByDescending » (Décroissant) permettent de trier une séquence sur plusieurs clé (propriétés). Nous pouvons avoir autant de clé « ThenBy » ou « ThenByDescending » que nous voulons. Cependant, ceux-ci ne sont utilisables que si un des opérateurs « OrderBy » ou « OrderByDescending » a été déclaré en premier lieu.

```
IEnumerable<Contact> QueryResult = Contacts  
    .OrderBy(c => c.AnneeDeNaissance)  
    .ThenByDescending(c => c.Nom);  
  
foreach (Contact c in QueryResult)  
{  
    Console.WriteLine("{0} {1}", c.Nom, c.AnneeDeNaissance);  
}
```

Expression de requête :

```
IEnumerable<Contact> QueryResult = from c in Contacts  
    orderby c.AnneeDeNaissance, c.Nom descending  
    select c;  
  
foreach (Contact c in QueryResult)  
{  
    Console.WriteLine("{0} {1}", c.Nom, c.AnneeDeNaissance);  
}
```

