Webh4ck partie V

© Stephane Rodriguez, 30 Avril 2002

Dans la suite logique des articles précédents, ce document détaille les aspects techniques permettant de mettre en œuvre un enregistreur d'actions de navigation web.

L'intérêt de l'enregistreur d'actions est comme son nom l'indique de pouvoir automatiser l'extraction de blocs de données web, et gagner en productivité.

La modélisation des actions va d'une exposition par API issue d'Internet Explorer à la génération d'un format spécifique basé XML évolutif.

1- API Internet Explorer

En résumé des articles précédents, on sait que les actions de navigation sur le web peuvent être classifiées en deux types : GET et POST. Et qu'indépendamment de cela, un serveur web et un navigateur web peuvent discuter de manière invisible par l'envoi et la réception d'entêtes spécifiques au protocole http connues principalement sous le nom de cookies.

L'API d'Internet Explorer permet à une application cliente d'être notifiée des paramètres GET et POST par l'appel de la méthode suivante :

[Extrait de la documentation MSDN]

virtual void OnBeforeNavigate2(LPCTSTR *lpszURL*, **DWORD** *nFlags*, **LPCTSTR** *lpszTargetFrameName*, **CByteArray&** *baPostedData*, **LPCTSTR** *lpszHeaders*, **BOOL*** *pbCancel* **);**

Parameters

lpszURL

Pointer to a string containing the URL to navigate to.

nFlags

Reserved for future use.

lpszTargetFrameName

A string that contains the name of the frame in which to display the resource, or **NULL** if no named frame is targeted for the resource.

baPostedData

A reference to a **CByteArray** object containing the data to send to the server if the HTTP POST transaction is being used.

lpszHeaders

A pointer to a string containing additional HTTP headers to send to the server (HTTP URLs only). The headers can specify such things as the action required of the server, the type of data being passed to the server, or a status code.

pbCancel

A pointer to a cancel flag. An application can set this parameter to nonzero to cancel the navigation operation, or to zero to allow it to proceed.

lpszURL contient l'action GET. Comme une action POST contient une action GET et des données supplémentaires, lpszURL a toujours du sens. Du reste, si l'URL était vide cela voudrait exactement dire que l'utilisateur n'a cliqué nulle part.

Exemple de valeur : lpszURL = "http://www.microsoft.com?default.asp&redir=50"

Les données spécifiques à la action POST sont passées dans **baPostedData**, et sont de la forme suivante :

redir=50
login=aaaaa
pwd=bbbbb

remarque : on constate que les paramètres POST pourraient être passés en GET et inversement. En fait l'action POST est une façon différente de faire…la même chose. Mais comme le web utilise intensivement les deux procédés, il faut les refléter.

2- Format de génération basé XML

L'action GET ou POST correspond à la transition entre deux pages web. Bien sûr, lorsque l'utilisateur tape une adresse URL au clavier, l'action GET correspondante permet d'accéder à une page : ce n'est pas en contradiction avec la phrase précédente.

Le format XML décrit donc un lien de transition.

Si l'on souhaite modéliser complètement la transition de A vers B, il est nécessaire d'exposer ce qu'est A et le moyen d'aller de A à B. De cette façon, la transition est complètement caractérisée et peut donner naissance à des applications très originales : en effet, la granularité est habituellement au niveau des pages puisque l'utilisateur accède à des pages web, or clairement une modélisation fine de la transition permet d'introduire la notion de composant web source et destination. Un composant web correspond à une zone d'intérêt pour l'utilisateur "normal", et à une source de données réutilisable pour l'utilisateur "expert".

Pour rester simple, il est possible de modéliser rapidement en XML l'équivalent des actions GET et POST, quitte à agrémenter cette modélisation par la suite d'un niveau de détails supplémentaire (autre article).

Les paramètres GET et POST jouent un rôle identique. Seule leur présentation étant différente, il suffit d'utiliser un élément XML de nom distinct pour chacun d'entre eux.

Exemples :

```
<get name="redir" value="50"/>
<post name="redir" value="50"/>
```

Lorsqu'il n'y a non pas un mais plusieurs paramètres, il suffit de les enchaîner :

```
<post name="redir" value="50"/>
<post name="login" value="aaaaa"/>
<post name="pwd" value="bbbbb"/>
```

Cette description doit être précédée de l'URL "principale" :

```
<url source="http://site.com/page1">
  <get name="redir" value="50"/>
</url>
```

L'élément url décrit la transition de la page courante à la page suivante. Pour décrire un ensemble de transitions, il suffit d'enchaîner les URLs :

```
<url source="http://site.com/page1"> <!—transition vers la page 2 -->
  <get name="redir" value="50"/>
</url>
<url source="http://site.com/page2"> <!—transition vers la page3  -->
  <post name="redir" value="50"/>
  <post name="login" value="aaaaa"/>
  <post name="pwd" value="bbbbb"/>
</url>
```

Jusqu'ici, la modélisation ne supporte que les transitions définies statiquement, numériquement, c'est-à-dire figées. Or l'évolution des sites web a entraîné une personnalisation de la navigation (les fameux MyYahoo, etc.), bref une dynamique qu'il s'agit de rendre compte dans la modélisation.
En particulier, entre deux pages, la valeur d'un paramètre doit pouvoir référencer un paramètre précédent, directement ou dans le cas général par l'exécution d'une formule associée.

Référencer un paramètre suppose conceptuellement qu'il y a au bout de la chaîne (par exemple en première page) de toute façon une valeur statique connue à l'avance. Ce n'est pas choquant qu'il n'y ait pas 100% de dynamique dans le modèle puisqu'en pratique le démarrage d'une navigation sur le web commence soit par la saisie au clavier d'une URL, le choix d'un favori ou le remplissage de données dans un formulaire. Dans les 3 cas, les valeurs sont statiques. Tout va bien.

Statistiquement les sites imposent souvent un login avant de pouvoir entrer dans le site à proprement parler. Les données du login sont autant de valeur statiques.

Dans l'exemple précédent, pour éviter de mettre deux fois la valeur 50, il est possible d'écrire :

```
<url source="http://site.com/page1"> <!—transition vers la page 2 -->
  <get name="redir" value="50"/>
</url>
<url source="http://site.com/page2"> <!—transition vers la page3 -->
  <post name="redir" value=":redir"/>
  <post name="login" value="aaaaa"/>
  <post name="pwd" value="bbbbb"/>
</url>
```

:redir est une syntaxe faisant partie du langage et qui indique que l'on fait référence à la valeur de ce paramètre. Si le paramètre n'existe pas réellement, ce qui correspond à une erreur de modélisation, la valeur retournée est une chaîne vide.

Pour pouvoir choisir la valeur de redir dans une transition plutôt qu'une autre, il est nécessaire de nommer celle-ci. Exemple :

```
<url name="page1" source="http://site.com/page1"> <!—transition vers la page 2 -->
  <get name="redir" value="50"/>
</url>
<url name="page2" source="http://site.com/page2"> <!—transition vers la page3 -->
  <post name="redir" value="page1:redir"/>
  <post name="login" value="aaaaa"/>
  <post name="pwd" value="bbbbb"/>
</url>
```

L'utilisation d'une formule permet de combiner une ou plusieurs valeurs. L'expérience montre qu'une formule comme la concaténation est très utilisée. Exemple :

```
<url name="page1" source="http://site.com/page1"> <!—transition vers la page 2 -->
  <get name="redir" value="50"/>
</url>
<url name="page2" source="http://site.com/page2"> <!—transition vers la page3 -->
  <post name="redir" value="concat('q=1&redir=',page1:redir,'&p=5')"/>
  <post name="login" value="aaaaa"/>
  <post name="pwd" value="bbbbb"/>
</url>
```

Ce qui va envoyer le paramètre post redir= q=1&redir=50&p=5

La formule utilisée en l'occurrence est **concat(arg1, arg2, ...)**

La transition entre deux pages doit non seulement tenir compte de ce que l'utilisateur fait, mais aussi de ce que le serveur web renvoie dans la transition. En pratique, il s'agit souvent des fameux cookies. Chaque cookie est une paire variable=valeur que le navigateur renvoie automatiquement dans toutes les actions se produisant ensuite.
Comme chaque action peut être à l'origine de plusieurs cookies, il est nécessaire dans la modélisation de pouvoir sélectionner un ou plusieurs cookies parmi ceux renvoyés. Un modèle de la forme **page1:redir[0]**, **page1:redir[1]** permet de répondre à cette demande. Les crochets entourent un indice démarrant à 0 et sélectionnant la ième occurrence nommée.

Cas d'utilisation : paramètre de session

Statistiquement, les sites hébergeant des serveurs comme Microsoft IIS utilisent par défaut un paramètre dit paramètre de session. C'est un identifiant généré par le serveur la première fois qu'on accède au site web en question, et qui suit l'utilisateur jusqu'à ce qu'il quitte le site. La paramètre de session permet de suivre son activité, particulièrement utile notamment pour lui permettre de faire des retours en arrière, etc. Il faut dire qu'avec la dynamique actuelle de génération des pages web on fait intervenir des extractions de bases de données, ces extractions étant forcément très limitées en taille elles définissent un contexte d'utilisation. Ce contexte est sauvegardé côté serveur, et le développeur du site peut le retrouver à tout moment à l'aide du paramètre de session. Sans contexte il n'y a pas de déterminisme des données présentées lors de la navigation, c'est donc très important.

Le paramètre de session est toutefois très volatile. Une fois le site quitté ou la machine éteinte, tout est perdu. C'est pourquoi les cookies sont si importants : ils permettent de stocker notamment le paramètre de session en local sur le disque dur de l'utilisateur dans un endroit privilégié. Et le navigateur web sait qu'à chaque fois qu'il accède au site il doit envoyer le cookie existant au domaine URL associé, d'où la possibilité de continuer une session, même au bout d'une semaine !
Les cookies n'ont pas que de bons effets, ils permettent insidieusement de suivre les utilisateurs à la trace sur un site, et même de croiser le profil de l'utilisateur sur plusieurs sites. Pourquoi ? parce que notamment via les bannières de publicité, hébergées par deux acteurs principaux dans le monde : DoubleClick (url de la forme http://ads.doubleclick/....) et RealMedia, ces derniers posent des cookies qui par définition sont renvoyés par le navigateur web à chaque fois que l'utilisateur obtient une bannière de publicité, pour ainsi dire à chaque fois qu'il surfe sur le web quel que soit le site.

3- Personnalisation de l'application tierce exécutant les scénarios générés

Proxy Port http : dans l'élément XML url, il est possible d'adjoindre l'attribut proxy dont la valeur est le nom d'une machine assurant le relais entre la machine de l'utilisateur et les sites web. Ce paramètre informatif est facultatif car dans de nombreux cas de figure l'application qui exécutera le scénario est dans un des cas de figure suivants :

- la connexion Internet est directe
- l'application exécutant le scénario se base sur IE, et notamment sur ses propriétés de connexion (cf panneau de configuration Windows) au sein desquels on trouve le proxy HTTP.
- l'application gère elle-même cette problématique

Les serveurs proxy sont souvent utilisés soit pour accéder à Internet de chez soi (le FAI fournit un proxy comme proxy.club-internet.fr), soit pour filtrer l'accès au web dans les entreprises. Sans l'indication de ce proxy, le scénario ne peut pas s'exécuter.
Exemple :

```
<url name="page1" source="http://www.microsoft.com" proxy="proxy.club-internet.fr"
method="GET">
</url>
```

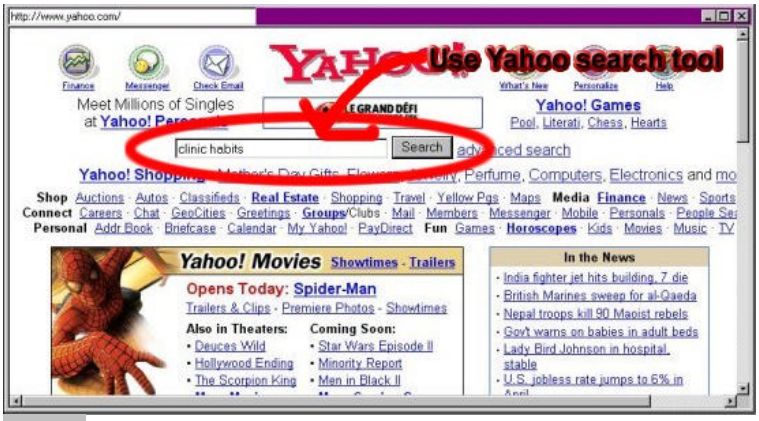
Dans le même ordre d'idée, il est possible d'indiquer un numéro de port http si la connexion au proxy se fait sur un autre port que 80. Exemple : `port="128"`.

Export Html : c'est la possibilité pour chaque page web de la stocker sur disque dur. C'est un souhait assez logique puisque l'exécution d'un scénario est bel et bien motivée par le fait d'accéder de manière automatique **à un contenu qui a de la valeur**. Et plutôt que de visionner forcément ce contenu lors de l'exécution de scénario, on peut souhaiter un archivage sur disque, ou envisager d'autres utilisations.
Exemple :

```
<url name="page1" source="http://www.microsoft.com" method="GET" output="page1.html">
</url>
```

4- Exemples complets (scénarios bruts)

- Yahoo, moteur de recherche



Page 1

Page 2

Scénario correspondant : (le scénario est le résultat d'un enregistrement, aucune modification n'a été effectuée pour améliorer sa lisibilité, ou l'optimiser)

```
<?xml version="1.0" encoding="UTF-8"?>
<webh4ck>
 <url name="page1" source="http://www.yahoo.com/" method="get">
  <header name="cookie" value="B=eikb244ub113n&b=2; Q=q1=AACAAAAAAAAAeg--&q2=PLFVaQ--"/>
 </url>
 <url name="page2" source="http://search.yahoo.com/bin/search?p=clinic+habits" method="get">
  <header name="cookie" value="B=eikb244ub113n&b=2; Q=q1=AACAAAAAAAAAeg--&q2=PLFVaQ--;
PU=t=1"/>
 </url>
</webh4ck>
```

Remarques : bien que le site Yahoo véhicule manifestement un objet session à travers un cookie, un test montre qu'il n'est pas obligatoire de passer le cookie. Ce n'est pas une règle, mais autant ne pas conserver des éléments facultatifs. Il n'est pas obligatoire de passer par la homepage non plus : une requête unique référencée par la page2 suffit à obtenir un résultat de recherche. Comme le paramètre est *url-encodé*, ce qui a été tapé au clavier **clinic habits** est devenu **clinic+habits**.

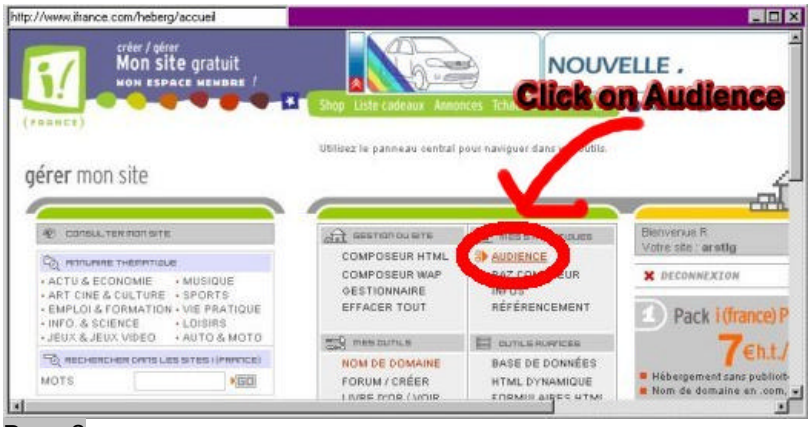
Automatisation obtenue après traitements manuels :

```
<?xml version="1.0" encoding="UTF-8"?>
<webh4ck>
<url name="page2" source="http://search.yahoo.com/bin/search?p=clinic+habits" method="get">
</url>
</webh4ck>
```

- IFrance : accéder aux statistiques de visite (page privée)

Page 1

Page 2



Page 3

Page 4

Scénario correspondant : (le scénario est le résultat d'un enregistrement, aucune modification n'a été effectuée pour améliorer sa lisibilité, ou l'optimiser)

```
<?xml version="1.0" encoding="UTF-8"?>
<webh4ck>
 <url name="page1" source="http://www.ifrance.com/heberg/" method="get">
  <header name="cookie" value="casto=1; I_RESOLUTION=1024"/>
 </url>
 <url name="page2" source="http://www.ifrance.com/heberg/accueil" method="get">
  <header name="cookie" value="casto=1; I_RESOLUTION=1024;
OPS_REFERER=http%3A//www.ifrance.com/heberg/"/>
 </url>
 <url name="page3" source="http://www.ifrance.com/heberg/accueil" method="post">
  <post name="Nom" value="arstlg"/>
  <post name="Pwd" value="XXXXXX"/>
  <post name="Produit" value="SITE"/>
  <post name="Back" value="http%3A%2F%2Fwww.ifrance.com%2Fheberg%2Faccueil.login"/>
```

```
  <post name="urlreloc" value="%2Fheberg%2Faccueil"/>
  <post name="ProfilId" value=""/>
  <post name="Page" value="AUTHENTIFICATION"/>
  <post name="Verif.x" value="25"/>
  <post name="Verif.y" value="9"/>
  <header name="cookie" value="casto=1; I_RESOLUTION=1024;
OPS_REFERER=http%3A//www.ifrance.com/heberg/; AuthId=2592913_MTMIQ; OPS_SITE_CUR_SITE=arstlg;
OPS_PROFIL_ID=91222EEHL"/>
 </url>
 <url name="page4" source="http://www.ifrance.com/stat?NomSite=arstlg" method="get">
  <header name="cookie" value="casto=1; I_RESOLUTION=1024;
OPS_REFERER=http%3A//www.ifrance.com/heberg/; AuthId=2592913_MTMIQ; OPS_SITE_CUR_SITE=arstlg;
OPS_PROFIL_ID=91222EEHL"/>
 </url>
</webh4ck>
```

Remarques : le site Ifrance passe par un formulaire de login. Comme pour de très nombreux sites, le login et le password passent en clair sur la ligne : la valeur XXXXXX n'est autre que le mot de passe. Tout individu sur la ligne peut récupérer ces informations et se faire passer pour l'utilisateur.
La page de démarrage saisie au clavier est http://www.ifrance.com, or ce n'est pas cette URL que l'on voit passer en page1. En fait le serveur ifrance a en l'occurrence procédé à un redirect, c'est-à-dire un aller-retour intermédiaire. Ce n'est pas important et n'empêche pas l'automatisation. C'est juste un résultat incohérent pour l'utilisateur.

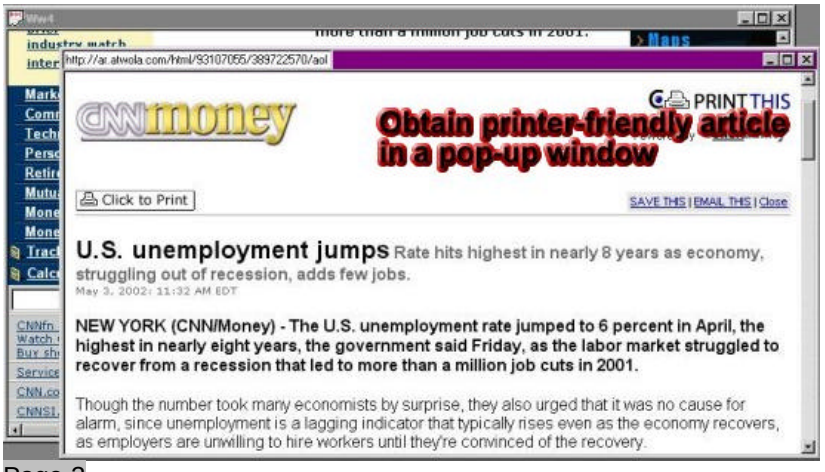
Automatisation obtenue après traitements manuels :

```
<?xml version="1.0" encoding="UTF-8"?>
<webh4ck>
 <url name="page2" source="http://www.ifrance.com/heberg/accueil" method="get">
  <header name="cookie" value="casto=1; I_RESOLUTION=1024;
OPS_REFERER=http%3A//www.ifrance.com/heberg/"/>
  <declare name="set-cookie"/>
 </url>
 <url name="page3" source="http://www.ifrance.com/heberg/accueil" method="post">
  <post name="Nom" value="arstlg"/>
  <post name="Pwd" value="XXXXXX"/>
  <post name="Produit" value="SITE"/>
  <post name="Back" value="http%3A%2F%2Fwww.ifrance.com%2Fheberg%2Faccueil.login"/>
  <post name="urlreloc" value="%2Fheberg%2Faccueil"/>
  <post name="ProfilId" value=""/>
  <post name="Page" value="AUTHENTIFICATION"/>
  <post name="Verif.x" value="25"/>
  <post name="Verif.y" value="9"/>
  <header name="cookie" value="page1:set-cookie[0]"/>
 </url>
 <url name="page4" source="http://www.ifrance.com/stat?NomSite=arstlg" method="get">
  <header name="cookie" value="page1:set-cookie[0]"/>
 </url>
</webh4ck>
```

- Site de news CNN : accéder à un article et l'imprimer

Page 1

Page 2



Page 3

Scénario correspondant : (le scénario est le résultat d'un enregistrement, aucune modification n'a été effectuée pour améliorer sa lisibilité, ou l'optimiser)

```
<?xml version="1.0" encoding="UTF-8"?>
<webh4ck>
 <url name="page1" source="" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page2" source="" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
```

```
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page3" source="" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page4"
source="http://ar.atwola.com/html/93103298/404269580/aol?SNM=HIDBF&CT=I&width=234&height=60&ta
rget=_top&TZ=-120" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page5" source="javascript:parent.adsFn1(93103298,234,60) http://www.cnn.com/"
method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page6" source="" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page7" source="" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page8" source="" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page9" source="" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page10" source="http://toolbar.netscape.com/tw_hat/iframe/cnn.html" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page11" source="" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page12" source="" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page13" source="" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page14" source="" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page15"
source="http://ar.atwola.com/html/93103309/404269580/aol?SNM=HIDBF&CT=I&width=88&height=31&tar
get=_top&TZ=-120" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page16"
source="http://ar.atwola.com/html/93114717/404269580/aol?SNM=HIDBF&CT=I&width=120&height=60&ta
rget=_top&TZ=-120" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page17"
source="http://ar.atwola.com/html/93103300/404269580/aol?SNM=HIDBF&CT=I&width=468&height=60&ta
rget=_top&TZ=-120" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page18" source="javascript:parent.adsFn1(93103309,88,31) http://www.cnn.com/"
method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page19" source="javascript:parent.adsFn1(93114717,120,60) http://www.cnn.com/"
```

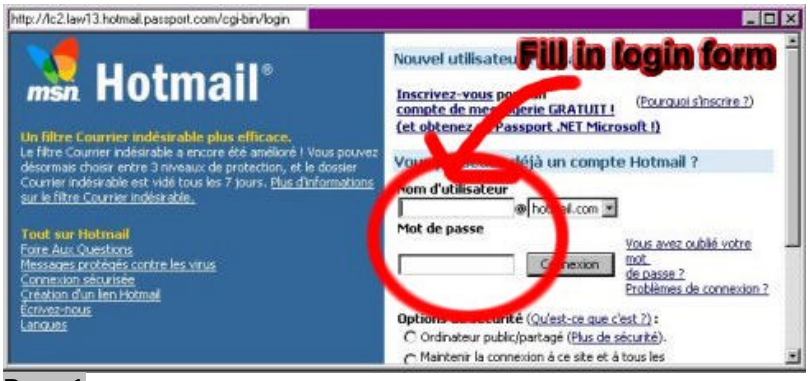
```
method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page20" source="javascript:parent.adsFn1(93103300,468,60) http://www.cnn.com/"
method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page21"
source="http://ar.atwola.com/html/93103299/404269580/aol?SNM=HIDBF&CT=I&width=120&height=60&ta
rget=_top&TZ=-120" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page22" source="javascript:parent.adsFn1(93103299,120,60) http://www.cnn.com/"
method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page23" source="http://www.cnn.com/" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page24" source="" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page25" source="" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page26" source="" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page27" source="" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page28" source="" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page29" source="http://toolbar.netscape.com/tw_hat/iframe/cnn.html" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page30"
source="http://ar.atwola.com/html/93114717/404311550/aol?SNM=HIDBF&CT=I&width=120&height=60&ta
rget=_top&TZ=-120" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page31"
source="javascript:parent.adsFn1(93114717,120,60) http://www.cnn.com/2002/TRAVEL/NEWS/05/03/c
leveland.evacuation/index.html" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page32"
source="http://ar.atwola.com/html/93102652/404311550/aol?SNM=HIDBF&CT=I&width=120&height=60&ta
rget=_top&TZ=-120" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page33"
source="javascript:parent.adsFn1(93102652,120,60) http://www.cnn.co m/2002/TRAVEL/NEWS/05/03/c
leveland.evacuation/index.html" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page34" source="" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page35" source="" method="get">
```

```
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page36"
source="http://ar.atwola.com/html/93102661/404311550/aol?SNM=HIDBF&CT=I&width=160&height=600&t
arget=_top&TZ=-120" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page37"
source="javascript:parent.adsFn1(93102661,160,600) http://www.cnn.com/2002/TRAVEL/NEWS/05/03/
cleveland.evacuation/index.html" method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page38"
source="http://www.cnn.com/2002/TRAVEL/NEWS/05/03/cleveland.evacuation/index.html"
method="get">
  <header name="cookie" value="EditionPopUp=seen(sh:1&id:0); CNNid=Gcf3013e9-4115088-
1020444435008-1; NGUserID=cf1947b9-3366-1020444376-6"/>
 </url>
 <url name="page39" source="" method="get">
  <header name="cookie" value="BIGipServerAPP=218802368.20480.000;
JServSessionIdprintthis=g3h815zn31.PT0; UID=261407474; x=wd985w2vW6lvs5vJod//vg==""/>
 </url>
 <url name="page40" source="" method="get">
  <header name="cookie" value="BIGipServerAPP=218802368.20480.000;
JServSessionIdprintthis=g3h815zn31.PT0; UID=261407474; x=wd985w2vW6lvs5vJod//vg==""/>
 </url>
 <url name="page41" source="" method="get">
  <header name="cookie" value="BIGipServerAPP=218802368.20480.000;
JServSessionIdprintthis=g3h815zn31.PT0; UID=261407474; x=wd985w2vW6lvs5vJod//vg==""/>
 </url>
 <url name="page42" source="" method="get">
  <header name="cookie" value="BIGipServerAPP=218802368.20480.000;
JServSessionIdprintthis=g3h815zn31.PT0; UID=261407474; x=wd985w2vW6lvs5vJod//vg==""/>
 </url>
 <url name="page43"
source="http://ar.atwola.com/html/93102707/404346090/aol?SNM=HIDBF&CT=I&width=468&height=60&ta
rget=_top&TZ=-120" method="get">
  <header name="cookie" value="BIGipServerAPP=218802368.20480.000;
JServSessionIdprintthis=g3h815zn31.PT0; UID=261407474; x=wd985w2vW6lvs5vJod//vg==""/>
 </url>
 <url name="page44"
source="javascript:parent.adsFn1(93102707,468,60) http://cnn.travel.printthis.clickability.co
m/pt/printThis?clickMap=printThis&fb=Y&url=http%3A//www.cnn.com/2002/TRAVEL/NEWS/05/03/clevela
nd.evacuation/index.html&title=CNN.com%20-
%20Security%20breach%20prompts%20airport%20evacuation%20-
%20May%203%2C%202002&random=0.7668914500472689&partnerID=2015&expire=-1" method="get">
  <header name="cookie" value="BIGipServerAPP=218802368.20480.000;
JServSessionIdprintthis=g3h815zn31.PT0; UID=261407474; x=wd985w2vW6lvs5vJod//vg==""/>
 </url>
 <url name="page45"
source="http://ar.atwola.com/html/93102707/404346090/aol?SNM=HIDBF&CT=I&width=468&height=60&ta
rget=_top&TZ=-120" method="get">
  <header name="cookie" value="BIGipServerAPP=218802368.20480.000;
JServSessionIdprintthis=g3h815zn31.PT0; UID=261407474; x=wd985w2vW6lvs5vJod//vg==""/>
 </url>
 <url name="page46"
source="javascript:parent.adsFn1(93102707,468,60) http://cnn.travel.printthis.clickability.co
m/pt/printThis?clickMap=printThis&fb=Y&url=http%3A//www.cnn.com/2002/TRAVEL/NEWS/05/03/clevela
nd.evacuation/index.html&title=CNN.com%20-
%20Security%20breach%20prompts%20airport%20evacuation%20-
%20May%203%2C%202002&random=0.7668914500472689&partnerID=2015&expire=-1" method="get">
  <header name="cookie" value="BIGipServerAPP=218802368.20480.000;
JServSessionIdprintthis=g3h815zn31.PT0; UID=261407474; x=wd985w2vW6lvs5vJod//vg==""/>
 </url>
 <url name="page47"
source="http://cnn.travel.printthis.clickability.com/pt/printThis?clickMap=printThis&fb=Y&url=
http%3A//www.cnn.com/2002/TRAVEL/NEWS/05/03/cleveland.evacuation/index.html&title=CNN.com%20-
%20Security%20breach%20prompts%20airport%20evacuation%20-
%20May%203%2C%202002&random=0.7668914500472689&partnerID=2015&expire=-1" method="get">
  <header name="cookie" value="BIGipServerAPP=218802368.20480.000;
JServSessionIdprintthis=g3h815zn31.PT0; UID=261407474; x=wd985w2vW6lvs5vJod//vg==""/>
 </url>
</webh4ck>
```

Remarques :

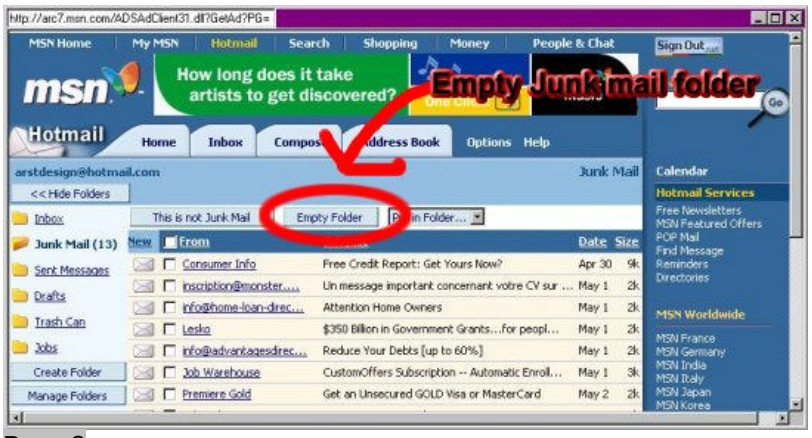
- comme pour Ifrance, l'url de démarrage http://www.cnn.com subit un redirect.
- il est très difficile de savoir dans ce scénario où se trouve l'action **Imprimer article**.
- le scénario comporte de nombreuses pseudo-actions, équivalent au total à 47 pages, alors que seules 3 pages du site ont été effectivement vues par l'utilisateur. En fait, le site CNN comme un grand nombre de sites provoque l'ouverture de multiples petites fenêtres rectangulaires à des fins publicitaires, d'où ces requêtes. En rouge dans le scénario sont relevées les 3 seules actions utiles (pages 4, 38 et 47).

- Hotmail : accéder au contenu de sa boite aux lettres, et vider les emails de spam



Page 1

Page 2



Page 3

Page 4

Scénario correspondant : (le scénario est le résultat d'un enregistrement, aucune modification n'a été effectuée pour améliorer sa lisibilité, ou l'optimiser)

```
<?xml version="1.0" encoding="UTF-8"?>
<webh4ck>
 <url name="page1" source="http://lc2.law13.hotmail.passport.com/cgi-bin/login" method="get">
  <header name="cookie" value="MSPDom=2; MSPPre=arstdesign@hotmail.com;
BrowserTest=Success%3f"/>
 </url>
 <url name="page2" source="https://loginnet.passport.com/ppsecure/post.srf" method="post">
  <post name="login" value="arstdesign"/>
  <post name="domain" value="hotmail.com"/>
  <post name="svc" value="mail"/>
  <post name="RemoteDAPost"
value="https%3A%2F%2Flogin.msnia.passport.com%2FFR%2Fppsecure%2Fpost.asp"/>
  <post name="passwd" value="XXXXXX"/>
  <post name="enter" value="Connexion"/>
  <post name="sec" value="no"/>
  <post name="curmbox" value="ACTIVE"/>
  <post name="js" value="yes"/>
  <post name="_lang" value="FR"/>
  <post name="beta" value=""/>
  <post name="ishotmail" value="1"/>
  <post name="mspp_shared" value=""/>
  <post name="id" value="2"/>
  <post name="fs" value="1"/>
  <post name="cb" value="_lang%253dFR"/>
  <post name="ct" value="1020444064"/>
  <post name="ru" value="http%3A%2F%2Fwww.hotmail.msn.com%2Fcgi-bin%2Fsbox"/>
  <header name="cookie" value="MSPDom=2; MSPPre=arstdesign@hotmail.com;
BrowserTest=Success%3f;
MSPSec=4i2Wbnu9XDASmw6vOkkIks4Gla9vfH2qXkD3cmG8VptA2pPn8mP6r2C7dKr2R7dv7L;
MSPAuth=4i2Wbnu9XDTaFf8ILqbNLm7Wla9vfH2qXk69E9RDMsdqrV52uWMFLxVdBEKOT!LCBJy!aIRJHf9lYR8O6QQSRn
rA$$;
MSPProf=4i2Wbnu9XHxARblSYyHPyK7zhhTCkk1I4NF7JI4mXOuGTQ8ptvLxKkj*c7ueTQzLVHS9hvZu2Ngts!gDMJlBqS
gA6tUkOhHHk7zGRAUFmD6NDOAv5ikIzqx94nwYPKXbrh81Sg9gl3DB2GsSVLTe2ReObOOX1WkmW5n8vh7H7nRcM$;
MSPVis=2"/>
 </url>
 <url name="page3" source="http://www.hotmail.msn.com/cgi-
bin/sbox?did=1&t=2i2Wbnu9XDvGt7lu77a!fWQ1s97jUSc8kyjdS7WnyiwU8FuR2*nNkNWEW4I9bpYvlNcYd417*fj40
6pGD8Fakpaw$$&p=2i2Wbnu9XBFhKpuZnbnqurb6Op3hDsvf0wiO8dSntLx!dG4T3t4r!Bn2ph!7Ix5UpGnU7bDlFxgpa0
8gu3ni53TpQ9c7gRLisADfKI6Buh1d0X13N1m*gpddekm2R29!gUy824*R1YqfUU1jobvhFI8ynUxL0iRP3meNiPQ1WAUh
!gNWJ7SQ6tyw$$&js=yes" method="get">
  <header name="cookie" value="lang=fr; MC1=V=2&GUID=8603b936a15e4397a44266d7a31eae3c;
MSPAuth=2i2Wbnu9XDvGt7lu77a%21fWQ1s97jUSc8kyjdS7WnyiwU8FuR2%2anNkNWEW4I9bpYvlNcYd417%2afj406pG
D8Fakpaw%24%24;
MSPProf=2i2Wbnu9XBFhKpuZnbnqurb6Op3hDsvf0wiO8dSntLx%21dG4T3t4r%21Bn2ph%217Ix5UpGnU7bDlFxgpa08g
u3ni53TpQ9c7gRLisADfKI6Buh1d0X13N1m%2agpddekm2R29%21gUy824%2aR1YqfUU1jobvhFI8ynUxL0iRP3meNiPQ1
WAUh%21gNWJ7SQ6tyw%24%24; HMP1=1;
HMSC0899=222arstdesign%40hotmail%2ecomtsENpEDkVS4pqIxvmF6kEXPopN0aXoI8wJPMAtzmVVRMgs73Ja48l%21
M%2aVY3GM1WQM1RQp52F%2aC6djtKZ%2alVBfWamvBhrDKoZcuyYmYRg6%21mK%21JWWhmb6v7RBmPBcG088q9SaSYSP4u
sQu0nsqgZVIe2AU5g%21YRgUh0%2aoZ9cJwrtBN1gkor%21hKEeeimdXM2M1rbLjUL2kM1%21dlSaFTSWF6FEnt4QN6eWk
I2EiGYoh%213LdGu%2arscRJgyktfdT6ICDR1ZzTcMbebvs5YkZn%2aTkKWQjpmleDs2Eq3jFUMLkN4HRdhR5REJLkt0yo
K2qEIuZ0%2171CZVIi8YhlISUkEanYDA%24%24"/>
 </url>
```

```
 <url name="page4" source="" method="get">
  <header name="cookie" value="lang=fr; MC1=V=2&GUID=8603b936a15e4397a44266d7a31eae3c;
MSPAuth=2i2Wbnu9XDvGt7lu77a%21fWQ1s97jUSc8kyjdS7WnyiwU8FuR2%2anNkNWEW4I9bpYvlNcYd417%2afj406pG
D8Fakpaw%24%24;
MSPProf=2i2Wbnu9XBFhKpuZnbnqurb6Op3hDsvf0wiO8dSntLx%21dG4T3t4r%21Bn2ph%217Ix5UpGnU7bDlFxgpa08g
u3ni53TpQ9c7gRLisADfKI6Buh1d0X13N1m%2agpddekm2R29%21gUy824%2aR1YqfUU1jobvhFI8ynUxL0iRP3meNiPQ1
WAUh%21gNWJ7SQ6tyw%24%24; HMP1=1;
HMSC0899=222arstdesign%40hotmail%2ecomtsENpEDkVS4pqIxvmF6kEXPopN0aXoI8wJPMAtzmVVRMgs73Ja48l%21
M%2aVY3GM1WQM1RQp52F%2aC6djtKZ%2alVBfWamvBhrDKoZcuyYmYRg6%21mK%21JWWhmb6v7RBmPBcG088q9SaSYSP4u
sQu0nsqgZVIe2AU5g%21YRgUh0%2aoZ9cJwrtBN1gkor%21hKEeeimdXM2M1rbLjUL2kM1%21dlSaFTSWF6FEnt4QN6eWk
I2EiGYoh%213LdGu%2arscRJgyktfdT6ICDR1ZzTcMbebvs5YkZn%2aTkKWQjpmleDs2Eq3jFUMLkN4HRdhR5REJLkt0yo
K2qEIuZ0%2171CZVIi8YhlISUkEanYDA%24%24"/>
 </url>
 <url name="page5" source="" method="get">
  <header name="cookie" value="lang=fr; MC1=V=2&GUID=8603b936a15e4397a44266d7a31eae3c;
MSPAuth=2i2Wbnu9XDvGt7lu77a%21fWQ1s97jUSc8kyjdS7WnyiwU8FuR2%2anNkNWEW4I9bpYvlNcYd417%2afj406pG
D8Fakpaw%24%24;
MSPProf=2i2Wbnu9XBFhKpuZnbnqurb6Op3hDsvf0wiO8dSntLx%21dG4T3t4r%21Bn2ph%217Ix5UpGnU7bDlFxgpa08g
u3ni53TpQ9c7gRLisADfKI6Buh1d0X13N1m%2agpddekm2R29%21gUy824%2aR1YqfUU1jobvhFI8ynUxL0iRP3meNiPQ1
WAUh%21gNWJ7SQ6tyw%24%24; HMP1=1;
HMSC0899=222arstdesign%40hotmail%2ecomtsENpEDkVS4pqIxvmF6kEXPopN0aXoI8wJPMAtzmVVRMgs73Ja48l%21
M%2aVY3GM1WQM1RQp52F%2aC6djtKZ%2alVBfWamvBhrDKoZcuyYmYRg6%21mK%21JWWhmb6v7RBmPBcG088q9SaSYSP4u
sQu0nsqgZVIe2AU5g%21YRgUh0%2aoZ9cJwrtBN1gkor%21hKEeeimdXM2M1rbLjUL2kM1%21dlSaFTSWF6FEnt4QN6eWk
I2EiGYoh%213LdGu%2arscRJgyktfdT6ICDR1ZzTcMbebvs5YkZn%2aTkKWQjpmleDs2Eq3jFUMLkN4HRdhR5REJLkt0yo
K2qEIuZ0%2171CZVIi8YhlISUkEanYDA%24%24"/>
 </url>
 <url name="page6"
source="http://popup.msn.com/lbpopupframe.asp?PG=HOTLBA&msid=000118E74C5C4644&country=DZ&UC=1"
method="get">
  <header name="cookie" value="lang=fr; MC1=V=2&GUID=8603b936a15e4397a44266d7a31eae3c;
MSPAuth=2i2Wbnu9XDvGt7lu77a%21fWQ1s97jUSc8kyjdS7WnyiwU8FuR2%2anNkNWEW4I9bpYvlNcYd417%2afj406pG
D8Fakpaw%24%24;
MSPProf=2i2Wbnu9XBFhKpuZnbnqurb6Op3hDsvf0wiO8dSntLx%21dG4T3t4r%21Bn2ph%217Ix5UpGnU7bDlFxgpa08g
u3ni53TpQ9c7gRLisADfKI6Buh1d0X13N1m%2agpddekm2R29%21gUy824%2aR1YqfUU1jobvhFI8ynUxL0iRP3meNiPQ1
WAUh%21gNWJ7SQ6tyw%24%24; HMP1=1;
HMSC0899=222arstdesign%40hotmail%2ecomtsENpEDkVS4pqIxvmF6kEXPopN0aXoI8wJPMAtzmVVRMgs73Ja48l%21
M%2aVY3GM1WQM1RQp52F%2aC6djtKZ%2alVBfWamvBhrDKoZcuyYmYRg6%21mK%21JWWhmb6v7RBmPBcG088q9SaSYSP4u
sQu0nsqgZVIe2AU5g%21YRgUh0%2aoZ9cJwrtBN1gkor%21hKEeeimdXM2M1rbLjUL2kM1%21dlSaFTSWF6FEnt4QN6eWk
I2EiGYoh%213LdGu%2arscRJgyktfdT6ICDR1ZzTcMbebvs5YkZn%2aTkKWQjpmleDs2Eq3jFUMLkN4HRdhR5REJLkt0yo
K2qEIuZ0%2171CZVIi8YhlISUkEanYDA%24%24"/>
 </url>
 <url name="page7"
source="http://arc5.msn.com/ADSAdClient31.dll?GetAd?PG=HOTENG?SC=PP?HM=043d2a5416524a0b233c441
42c0545760a5118484754521e344b062d29?LOC=H?TF=adframe?PUID=000118E74C5C4644?UC=1" method="get">
  <header name="cookie" value="lang=fr; MC1=V=2&GUID=8603b936a15e4397a44266d7a31eae3c;
MSPAuth=2i2Wbnu9XDvGt7lu77a%21fWQ1s97jUSc8kyjdS7WnyiwU8FuR2%2anNkNWEW4I9bpYvlNcYd417%2afj406pG
D8Fakpaw%24%24;
MSPProf=2i2Wbnu9XBFhKpuZnbnqurb6Op3hDsvf0wiO8dSntLx%21dG4T3t4r%21Bn2ph%217Ix5UpGnU7bDlFxgpa08g
u3ni53TpQ9c7gRLisADfKI6Buh1d0X13N1m%2agpddekm2R29%21gUy824%2aR1YqfUU1jobvhFI8ynUxL0iRP3meNiPQ1
WAUh%21gNWJ7SQ6tyw%24%24; HMP1=1;
HMSC0899=222arstdesign%40hotmail%2ecomtsENpEDkVS4pqIxvmF6kEXPopN0aXoI8wJPMAtzmVVRMgs73Ja48l%21
M%2aVY3GM1WQM1RQp52F%2aC6djtKZ%2alVBfWamvBhrDKoZcuyYmYRg6%21mK%21JWWhmb6v7RBmPBcG088q9SaSYSP4u
sQu0nsqgZVIe2AU5g%21YRgUh0%2aoZ9cJwrtBN1gkor%21hKEeeimdXM2M1rbLjUL2kM1%21dlSaFTSWF6FEnt4QN6eWk
I2EiGYoh%213LdGu%2arscRJgyktfdT6ICDR1ZzTcMbebvs5YkZn%2aTkKWQjpmleDs2Eq3jFUMLkN4HRdhR5REJLkt0yo
K2qEIuZ0%2171CZVIi8YhlISUkEanYDA%24%24"/>
 </url>
 <url name="page8" source="http://lw7fd.law7.hotmail.msn.com/cgi-
bin/hmhome?curmbox=F000000001&a=7481be0ea1c4f7990dad9c992440defb&fti=yes&_lang=EN"
method="get">
  <header name="cookie" value="lang=fr; MC1=V=2&GUID=8603b936a15e4397a44266d7a31eae3c;
MSPAuth=2i2Wbnu9XDvGt7lu77a%21fWQ1s97jUSc8kyjdS7WnyiwU8FuR2%2anNkNWEW4I9bpYvlNcYd417%2afj406pG
D8Fakpaw%24%24;
MSPProf=2i2Wbnu9XBFhKpuZnbnqurb6Op3hDsvf0wiO8dSntLx%21dG4T3t4r%21Bn2ph%217Ix5UpGnU7bDlFxgpa08g
u3ni53TpQ9c7gRLisADfKI6Buh1d0X13N1m%2agpddekm2R29%21gUy824%2aR1YqfUU1jobvhFI8ynUxL0iRP3meNiPQ1
WAUh%21gNWJ7SQ6tyw%24%24; HMP1=1;
HMSC0899=222arstdesign%40hotmail%2ecomtsENpEDkVS4pqIxvmF6kEXPopN0aXoI8wJPMAtzmVVRMgs73Ja48l%21
M%2aVY3GM1WQM1RQp52F%2aC6djtKZ%2alVBfWamvBhrDKoZcuyYmYRg6%21mK%21JWWhmb6v7RBmPBcG088q9SaSYSP4u
sQu0nsqgZVIe2AU5g%21YRgUh0%2aoZ9cJwrtBN1gkor%21hKEeeimdXM2M1rbLjUL2kM1%21dlSaFTSWF6FEnt4QN6eWk
I2EiGYoh%213LdGu%2arscRJgyktfdT6ICDR1ZzTcMbebvs5YkZn%2aTkKWQjpmleDs2Eq3jFUMLkN4HRdhR5REJLkt0yo
K2qEIuZ0%2171CZVIi8YhlISUkEanYDA%24%24"/>
 </url>
 <url name="page9" source="" method="get">
  <header name="cookie" value="lang=fr; MC1=V=2&GUID=8603b936a15e4397a44266d7a31eae3c;
MSPAuth=2i2Wbnu9XDvGt7lu77a%21fWQ1s97jUSc8kyjdS7WnyiwU8FuR2%2anNkNWEW4I9bpYvlNcYd417%2afj406pG
D8Fakpaw%24%24;
MSPProf=2i2Wbnu9XBFhKpuZnbnqurb6Op3hDsvf0wiO8dSntLx%21dG4T3t4r%21Bn2ph%217Ix5UpGnU7bDlFxgpa08g
u3ni53TpQ9c7gRLisADfKI6Buh1d0X13N1m%2agpddekm2R29%21gUy824%2aR1YqfUU1jobvhFI8ynUxL0iRP3meNiPQ1
```

```
WAUh%21gNWJ7SQ6tyw%24%24; HMP1=1;
HMSC0899=222arstdesign%40hotmail%2ecomtsENpEDkVS4pqIxvmF6kEXPopN0aXoI8wJPMAtzmVVRMgs73Ja48l%21
M%2aVY3GM1WQM1RQp52F%2aC6djtKZ%2alVBfWamvBhrDKoZcuyYmYRg6%21mK%21JWWhmb6v7RBmPBcG088q9SaSYSP4u
sQu0nsqgZVIe2AU5g%21YRgUh0%2aoZ9cJwrtBN1gkor%21hKEeeimdXM2M1rbLjUL2kM1%21dlSaFTSWF6FEnt4QN6eWk
I2EiGYoh%213LdGu%2arscRJgyktfdT6ICDR1ZzTcMbebvs5YkZn%2aTkKWQjpmleDs2Eq3jFUMLkN4HRdhR5REJLkt0yo
K2qEIuZ0%2171CZVIi8YxlISUkEa3YDA%24%24"/>
 </url>
 <url name="page10"
source="http://arc7.msn.com/ADSAdClient31.dll?GetAd?PG=HOTBOS?SC=LG?HM=043d2a5416524a0b233c441
42c0545760a5118484754521e344b062d29?LOC=I?TF=adframe?PUID=000118E74C5C4644?UC=1" method="get">
  <header name="cookie" value="lang=fr; MC1=V=2&GUID=8603b936a15e4397a44266d7a31eae3c;
MSPAuth=2i2Wbnu9XDvGt7lu77a%21fWQ1s97jUSc8kyjdS7WnyiwU8FuR2%2anNkNWEW4I9bpYvlNcYd417%2afj406pG
D8Fakpaw%24%24;
MSPProf=2i2Wbnu9XBFhKpuZnbnqurb6Op3hDsvf0wiO8dSntLx%21dG4T3t4r%21Bn2ph%217Ix5UpGnU7bDlFxgpa08g
u3ni53TpQ9c7gRLisADfKI6Buh1d0X13N1m%2agpddekm2R29%21gUy824%2aR1YqfUU1jobvhFI8ynUxL0iRP3meNiPQ1
WAUh%21gNWJ7SQ6tyw%24%24; HMP1=1;
HMSC0899=222arstdesign%40hotmail%2ecomtsENpEDkVS4pqIxvmF6kEXPopN0aXoI8wJPMAtzmVVRMgs73Ja48l%21
M%2aVY3GM1WQM1RQp52F%2aC6djtKZ%2alVBfWamvBhrDKoZcuyYmYRg6%21mK%21JWWhmb6v7RBmPBcG088q9SaSYSP4u
sQu0nsqgZVIe2AU5g%21YRgUh0%2aoZ9cJwrtBN1gkor%21hKEeeimdXM2M1rbLjUL2kM1%21dlSaFTSWF6FEnt4QN6eWk
I2EiGYoh%213LdGu%2arscRJgyktfdT6ICDR1ZzTcMbebvs5YkZn%2aTkKWQjpmleDs2Eq3jFUMLkN4HRdhR5REJLkt0yo
K2qEIuZ0%2171CZVIi8YxlISUkEa3YDA%24%24"/>
 </url>
 <url name="page11" source="http://lw7fd.law7.hotmail.msn.com/cgi-
bin/HoTMaiL?curmbox=F000000005&chkprotector=1&curmbox=F000000005&a=2918357cece813a71e4c57edc9e
35836" method="get">
  <header name="cookie" value="lang=fr; MC1=V=2&GUID=8603b936a15e4397a44266d7a31eae3c;
MSPAuth=2i2Wbnu9XDvGt7lu77a%21fWQ1s97jUSc8kyjdS7WnyiwU8FuR2%2anNkNWEW4I9bpYvlNcYd417%2afj406pG
D8Fakpaw%24%24;
MSPProf=2i2Wbnu9XBFhKpuZnbnqurb6Op3hDsvf0wiO8dSntLx%21dG4T3t4r%21Bn2ph%217Ix5UpGnU7bDlFxgpa08g
u3ni53TpQ9c7gRLisADfKI6Buh1d0X13N1m%2agpddekm2R29%21gUy824%2aR1YqfUU1jobvhFI8ynUxL0iRP3meNiPQ1
WAUh%21gNWJ7SQ6tyw%24%24; HMP1=1;
HMSC0899=222arstdesign%40hotmail%2ecomtsENpEDkVS4pqIxvmF6kEXPopN0aXoI8wJPMAtzmVVRMgs73Ja48l%21
M%2aVY3GM1WQM1RQp52F%2aC6djtKZ%2alVBfWamvBhrDKoZcuyYmYRg6%21mK%21JWWhmb6v7RBmPBcG088q9SaSYSP4u
sQu0nsqgZVIe2AU5g%21YRgUh0%2aoZ9cJwrtBN1gkor%21hKEeeimdXM2M1rbLjUL2kM1%21dlSaFTSWF6FEnt4QN6eWk
I2EiGYoh%213LdGu%2arscRJgyktfdT6ICDR1ZzTcMbebvs5YkZn%2aTkKWQjpmleDs2Eq3jFUMLkN4HRdhR5REJLkt0yo
K2qEIuZ0%2171CZVIi8YxlISUkEa3YDA%24%24"/>
 </url>
 <url name="page12" source="" method="get">
  <header name="cookie" value="lang=fr; MC1=V=2&GUID=8603b936a15e4397a44266d7a31eae3c;
MSPAuth=2i2Wbnu9XDvGt7lu77a%21fWQ1s97jUSc8kyjdS7WnyiwU8FuR2%2anNkNWEW4I9bpYvlNcYd417%2afj406pG
D8Fakpaw%24%24;
MSPProf=2i2Wbnu9XBFhKpuZnbnqurb6Op3hDsvf0wiO8dSntLx%21dG4T3t4r%21Bn2ph%217Ix5UpGnU7bDlFxgpa08g
u3ni53TpQ9c7gRLisADfKI6Buh1d0X13N1m%2agpddekm2R29%21gUy824%2aR1YqfUU1jobvhFI8ynUxL0iRP3meNiPQ1
WAUh%21gNWJ7SQ6tyw%24%24; HMP1=1;
HMSC0899=222arstdesign%40hotmail%2ecomtsENpEDkVS4pqIxvmF6kEXPopN0aXoI8wJPMAtzmVVRMgs73Ja48l%21
M%2aVY3GM1WQM1RQp52F%2aC6djtKZ%2alVBfWamvBhrDKoZcuyYmYRg6%21mK%21JWWhmb6v7RBmPBcG088q9SaSYSP4u
sQu0nsqgZVIe2AU5g%21YRgUh0%2aoZ9cJwrtBN1gkor%21hKEeeimdXM2M1rbLjUL2kM1%21dlSaFTSWF6FEnt4QN6eWk
I2EiGYoh%213LdGu%2arscRJgyktfdT6ICDR1ZzTcMbebvs5YkZn%2aTkKWQjpmleDs2Eq3jFUMLkN4HRdhR5REJLkt0yo
K2qEIuZ0%2171CZVIi8YxlISUkEa3YDA%24%24"/>
 </url>
 <url name="page13"
source="http://arc7.msn.com/ADSAdClient31.dll?GetAd?PG=HOTBOS?SC=LG?HM=043d2a5416524a0b233c441
42c0545760a5118484754521e344b062d29?LOC=I?TF=adframe?PUID=000118E74C5C4644?UC=1" method="get">
  <header name="cookie" value="lang=fr; MC1=V=2&GUID=8603b936a15e4397a44266d7a31eae3c;
MSPAuth=2i2Wbnu9XDvGt7lu77a%21fWQ1s97jUSc8kyjdS7WnyiwU8FuR2%2anNkNWEW4I9bpYvlNcYd417%2afj406pG
D8Fakpaw%24%24;
MSPProf=2i2Wbnu9XBFhKpuZnbnqurb6Op3hDsvf0wiO8dSntLx%21dG4T3t4r%21Bn2ph%217Ix5UpGnU7bDlFxgpa08g
u3ni53TpQ9c7gRLisADfKI6Buh1d0X13N1m%2agpddekm2R29%21gUy824%2aR1YqfUU1jobvhFI8ynUxL0iRP3meNiPQ1
WAUh%21gNWJ7SQ6tyw%24%24; HMP1=1;
HMSC0899=222arstdesign%40hotmail%2ecomtsENpEDkVS4pqIxvmF6kEXPopN0aXoI8wJPMAtzmVVRMgs73Ja48l%21
M%2aVY3GM1WQM1RQp52F%2aC6djtKZ%2alVBfWamvBhrDKoZcuyYmYRg6%21mK%21JWWhmb6v7RBmPBcG088q9SaSYSP4u
sQu0nsqgZVIe2AU5g%21YRgUh0%2aoZ9cJwrtBN1gkor%21hKEeeimdXM2M1rbLjUL2kM1%21dlSaFTSWF6FEnt4QN6eWk
I2EiGYoh%213LdGu%2arscRJgyktfdT6ICDR1ZzTcMbebvs5YkZn%2aTkKWQjpmleDs2Eq3jFUMLkN4HRdhR5REJLkt0yo
K2qEIuZ0%2171CZVIi8YxlISUkEa3YDA%24%24"/>
 </url>
 <url name="page14" source="http://lw7fd.law7.hotmail.msn.com/cgi-
bin/HoTMaiL?curmbox=F000000005&a=2918357cece813a71e4c57edc9e35836&DoEmpty=1" method="get">
  <header name="cookie" value="lang=fr; MC1=V=2&GUID=8603b936a15e4397a44266d7a31eae3c;
MSPAuth=2i2Wbnu9XDvGt7lu77a%21fWQ1s97jUSc8kyjdS7WnyiwU8FuR2%2anNkNWEW4I9bpYvlNcYd417%2afj406pG
D8Fakpaw%24%24;
MSPProf=2i2Wbnu9XBFhKpuZnbnqurb6Op3hDsvf0wiO8dSntLx%21dG4T3t4r%21Bn2ph%217Ix5UpGnU7bDlFxgpa08g
u3ni53TpQ9c7gRLisADfKI6Buh1d0X13N1m%2agpddekm2R29%21gUy824%2aR1YqfUU1jobvhFI8ynUxL0iRP3meNiPQ1
WAUh%21gNWJ7SQ6tyw%24%24; HMP1=1;
HMSC0899=222arstdesign%40hotmail%2ecomtsENpEDkVS4pqIxvmF6kEXPopN0aXoI8wJPMAtzmVVRMgs73Ja48l%21
M%2aVY3GM1WQM1RQp52F%2aC6djtKZ%2alVBfWamvBhrDKoZcuyYmYRg6%21mK%21JWWhmb6v7RBmPBcG088q9SaSYSP4u
sQu0nsqgZVIe2AU5g%21YRgUh0%2aoZ9cJwrtBN1gkor%21hKEeeimdXM2M1rbLjUL2kM1%21dlSaFTSWF6FEnt4QN6eWk
I2EiGYoh%213LdGu%2arscRJgyktfdT6ICDR1ZzTcMbebvs5YkZn%2aTkKWQjpmleDs2Eq3jFUMLkN4HRdhR5REJLkt0yo
```

```
K2qEIuZ0%2171CZVIi8YxlISUkEa3YDA%24%24"/>
 </url>
</webh4ck>
```

Remarques :

- la page de démarrage http://www.hotmail.com subit un redirect.
- alors que l'utilisateur ne fait que remplir son formulaire d'authentification suivi de deux clicks, le scénario ne répercute pas moins de 14 pages. En fait à l'instar de CNN, Hotmail fait de nombreuses requêtes "inutiles" pour l'utilisateur. Mis à part les soucis de communication publicitaire, la réalité est que le contenu de la page Hotmail tente de lancer le composant MSN Messenger (au nez et à la barbe de l'utilisateur) pour qu'il synchronise son contenu. Ceci a un effet direct sur le scénario, et est subi par l'utilisateur.
- le souci de sécurité fait prendre aux responsables de Hotmail des mesures quasi grotesques quant au passage des paramètres de session et autres authentifications : si quelqu'un est capable de voir passer 10 caractères dans un cookie, il est certainement tout aussi capable de voir passer 100 caractères. Il n'y a donc pas plus de sécurité qu'ailleurs.

Remarques globales :

- sur les 4 sites, deux sont accessibles via une bannière de login. Dans les deux cas de figure, des paramètres de session sont générés dynamiquement. L'enregistrement des actions les voit passer, mais il est évident que ces scénarios ne sont pas réutilisables de nouveau sans modifications manuelles et mise sous forme de variables du ou des cookies.
- le scénario de Yahoo est rejouable tel quel. Celui de CNN aussi. Preuve que les sites "publics" se laissent automatiser sans autre forme de procès.
- les scénarios enregistrés sont finalement assez verbeux, et subissent malheureusement les assauts des ouvertures intempestives des fenêtres popups. Pour supprimer automatiquement ces pseudo-actions, il est nécessaire de créer des règles. Par ailleurs pour ces pseudo-actions il est parfois possible à l'enregistreur de scénario de ne pas en tenir compte : trois cas de figure : l'attribut source de l'url est vide ou contient un appel à un code javascript ; l'action est exactement répétée plus d'une fois ; il y a plus d'une action enregistrée par clic (pas forcément évident à synchroniser).

5- Modélisation de la destination

Dans la modélisation de la transition de A vers B, B n'est pas forcément une page. C'est essentiellement une page mais des besoins spécifiques peuvent requérir un niveau de détail plus fin, un sous-ensemble de la page web B.
En pratique, on a besoin de celle-ci pour extraire en cours de scénario non pas des pages web entières, mais certaines parties privilégiées.

La modélisation risque d'être assez originale puisque l'utilisateur accède à une page web et fait lui-même l'effort de regarder et lire dans celle-ci les rubriques ou composants qui l'intéressent. Cet effort est intellectuel, il n'y a pas d'action manuelle.

La modélisation est d'autant plus ardue qu'il n'y a pas de standard de conception de page web (et cela se comprend, sinon le web serait bien terne). Chaque site fait son propre design. Il faut dire que les besoins d'un site éditorial ne sont pas les mêmes qu'un site de prestation de services. Et ce même si objectivement on trouve de manière quasi systématique une barre de navigation dans les rubriques et un contenu dans l'espace principal de la page.
Comme ces pages sont faites avec html, et que ce protocole n'a qu'une sémantique de présentation, et non de description de contenu, il est clairement impossible de supposer que les composants de la page respectent une nomenclature : barre de navigation principale, barre de liens d'enrichissement, barre corporate, …

Le seul standard qu'il y ait c'est le code html lui-même.

En fait il n'y a pas de notion de composant web à proprement parler. Il se trouve que statistiquement les barres sont clairement identifiables entre elles grâce à l'utilisation de tableaux englobants, moyen de séparer le contenu à coup sûr puisque tout le code HTML non déclaré à l'intérieur d'un tableau sera affiché…à l'extérieur.

Pour identifier un composant, il est donc possible dans certains cas de modéliser la position d'un tableau dans le code HTML, soit par un nom qu'il porterait dans un attribut, soit via l'utilisation d'une ligne de commentaire additionnelle, soit par la concordance avec une suite de tags HTML très particuliers qui le singulariserait, soit encore parce que le tableau, comme tout tag HTML, se trouve dans une arborescence dite DOM et que tout nœud de l'arbre est accessible.

Cas du nommage par attribut :

```
<table myname="tableau1" width=300 height=200>
…
</table>
```

Cas du nommage via ligne de commentaire :

```
<!—BEGIN tableau1 -->
<table width=300 height=200>
…
</table>
<!—END tableau1 -->
```

Cas de motif reconnaissable de tags HTML :

```
<table width=300 height=200 border=0>
 <tr rowspan=4>
…
  </tr>
</table>
```

On peut imaginer que les attributs border et rowspan ne soient utilisés que pour ce tableau dans la totalité de la page web, ou que les valeurs associées ne soient telles que pour ce tableau.

Cas de l'arborescence DOM :

```
<html>
 <head>
   <body>
       <table …>
       <table myname="tableau1"…>
```

Le tag table qui nous intéresse est le deuxième fils du tag body, lui-même fils de head, lui-même fils de html.

Dans tous les cas de figure, il est clair que l'identification d'un composant par une machine passe par la déclaration d'un "pattern matching" c'est-à-dire d'une logique de relations entre des tags HTML. Cette logique fait elle-même partie d'une autre modélisation basée soit sur un standard comme Xpath (http://www.w3c.org) soit sur une modélisation ad hoc.

Que devient ce pattern matching lorsque le contenu du site change ? et a priori le contenu du site change tous les jours, c'est donc particulièrement critique ! La première chose à dire est que, puisque la modélisation est basée sur un simple enregistrement de navigation, elle ne peut se prévaloir d'être générique – impossible d'utiliser un scénario d'un site pour le compte d'un autre -, ce qui semble assez facile à comprendre, mais par contre il est vrai que plus la destination ciblée est précise, plus elle est sensible aux moindres changements. Potentiellement, le scénario ne fonctionne plus. Il faut soit le refaire entièrement, soit l'ajuster. Il n'est pas déraisonnable que l'ajustement puisse se faire de manière automatique.
Ces changements dans le contenu sont de plusieurs types :

• le contenu a réellement totalement changé, au sens où la présentation a totalement changé. Très rare statistiquement, cela correspond à une refonte d'un site.
• le contenu n'a pas changé mais la présentation est totalement différente : on est dans le pire cas de figure : le site a une très forte composante personnalisation ou une très forte propension à présenter sans cesse différemment un même message. Pas d'automatisation possible a priori, sauf si les composants restent identifiables dans le code HTML. Cas très rare malgré tout, car l'identité visuelle voulue par les sites incite à ne pas présenter les contenus n'importe comment en permanence.
• le contenu a changé mais le taggage HTML (motifs) est identique. C'est le cas typique d'un site de news avec un composant de news en bref : le message change en permanence (exemple des dépêches AFP) mais la présentation reste inchangée. Clairement ceci est compatible et même favorable avec l'automatisation.
• le contenu a changé et la présentation a changé un peu. En particulier certains styles HTML sont changés. Dans ce cas de deux choses l'une, soit le pattern matching est suffisamment indépendant des styles HTML, soit l'algorithme qui exécute le pattern matching est relativement insensible à ces changements. Dans tous les cas, l'automatisation est possible et la maintenance des scénarios potentiellement nulle.
• le contenu a changé, pas significativement visuellement (voire pas du tout), mais le taggage HTML a totalement changé. Là il faut plutôt exploiter un pattern matching nominatif sur les composants, sinon il est certain ou quasi-certain que l'automatisation n'a pas de sens.

En résumé de cette discussion sur le ciblage précis des composants d'une page web, il apparaît que plus le ciblage est précis, plus le coût de maintenance risque d'être élevé, sauf si la modélisation pattern matching est suffisamment bien pensée, le site se prête plutôt facilement à l'extraction de données, et que l'extraction des données à proprement parler est suffisamment robuste pour passer à travers des modifications de styles (et non des modifications de structure).

6- Modélisation de la source

De la même manière que la destination peut être précisément modélisée, la source peut l'être également. La source correspond à ce sur quoi l'utilisateur clique, c'est donc essentiellement un lien hypertexte click.

En pratique il y a plusieurs possibilités :
• l'utilisateur vient de saisir une adresse URL complète, c'est un cas de figure où il n'y a pas de source. Mot-clef = **newurl**.
• l'utilisateur vient de cliquer sur un lien, la source est un lien hypertexte sur un texte, une image, … Mot-clef = **link, text, DOMpath= link, bmp, DOMpath=**
• l'utilisateur vient de cliquer sur le bouton d'envoi d'un formulaire, la source est un élément de formulaire (en général le formulaire est d'un seul tenant et n'a qu'une seule vocation). Mot-clef = **link, form, DOMpath=**
• l'action provient de l'exécution d'un code javascript, comme par exemple location.href="http://www.monsite.com". La source est difficilement modélisable à part que l'on peut savoir qu'il s'agit de javascript. Mot-clef = **javascript**.
• l'action provient d'une redirection serveur, c'est-à-dire d'un aller-retour automatique. Mot-clef = **redirect**.
• l'utilisateur a cliqué sur un lien tout en demandant l'ouverture d'une nouvelle fenêtre. Mot-clef supplémentaire = **newwindow**.

DOMPath est un paramètre identifiant un tag HTML dans l'arborescence du code source HTML.

7- Coût de la maintenance de la modélisation

Comme toute modélisation, il y a un prix à payer. Il est clair que les actions GET/POST n'ont de valeur qu'à un instant t. La grande hypothèse faite, et qui motive le modèle, c'est que les sites ne changent pas souvent leur construction, seulement une fois tous les 6 mois par exemple, ce qui permet de voir venir.
Bien sûr, il n'en va pas de même du contenu d'une page web, qui lui change tous les jours. Les actions GET/POST sont indépendantes du contenu à proprement parler, ce qui est une bonne nouvelle!

Mieux que cela encore, c'est que si l'outil d'enregistrement des actions est totalement intuitif, il n'est pas lassant ni rébarbatif de mettre à jour un enregistrement. On peut comprendre toutefois que dans le cas d'un environnement de production idéal, il n'est pas constructif de refaire une opération faite quelque temps auparavant.

Si le site change complètement sa construction, clairement l'enregistrement devra être refait. Par contre, si le site ne change qu'en surface, il n'est pas impossible que :

- les actions GET/POST continuent à fonctionner normalement, alors même que l'apparence visuelle du site a changé. C'est par exemple le cas de Yahoo et de son moteur de recherche : un logiciel automatisé écrit il y a deux ans fonctionne toujours.
- le logiciel enregistreur essaye de concorder le scénario connu d'actions avec la réalité des actions au moment de l'exécution du scénario. En clair, il n'est pas impossible d'essayer de coller des morceaux disloqués et de suivre des actions GET/POST menant au but désiré. Le logiciel enregistreur peut alors mettre à jour automatiquement le scénario. Tout aura été fait de manière transparente.
- une modification manuelle du scénario peut être requise. Il est nécessaire à l'utilisateur d'avoir lu cet article et d'être capable de faire des modifications "dans le code". Ce dernier cas est évidemment le plus douloureux et mérite à peine d'être considéré. En effet, mieux vaut réenregistrer un scénario complet, ce qui ne nécessite aucune compétence particulière.

8- Ajout automatique de la dynamique d'un scénario

Il ne faut pas se voiler la face, l'enregistreur d'actions est très bête. Il met au format XML les données qu'il voit passer. Dans le même temps, beaucoup de sites utilisent les allers-retours de paramètres pour des besoins d'identification et de session, ce qui entraîne une dynamique inévitable. Or la dynamique dans le modèle du scénario suppose non seulement une connaissance extrêmement fine des actions en question, mais également le fait d'aller changer le format XML soi-même.
Ce n'est pas viable.
Il y a vraiment un très fort intérêt à ce que l'enregistreur d'actions soit capable, de lui-même, de proche en proche de se rendre compte que le passage dynamique de certaines valeurs est nécessaire et de faire lui-même ce passage.
On peut difficilement en vouloir à l'enregistreur d'actions de ne pas être capable simplement de déduire qu'il faut à un certain endroit concaténer dynamiquement plusieurs valeurs pour en faire une autre, encore que. Mais il semble possible techniquement de "voir passer" réellement des paramètres type session, c'est-à-dire de distinguer automatiquement les paramètres statiques des paramètres dynamiques. Et c'est cela le principal.
Comment faire ? Un principe simple, lorsque l'utilisateur enregistre un scénario, il n'est pas interdit d'exécuter ce dernier à travers deux instances distinctes de scénario, le but étant précisément d'avoir deux sessions serveurs distinctes. Alors même que l'utilisateur ne va enregistrer qu'une fois son scénario. Un scénario physique et visible, pour deux scénarios logiques et non visibles. Ces scénarios ont une particularité très intéressante : ils sont comparables, ce qui permet de faire ressortir les valeurs dynamiques, et donc les identifiants de session.
Dans un monde idéal, tout ceci est transparent pour l'utilisateur. Mais on peut en fait imaginer dans une logique de développement de code incrémentale que l'enregistreur publie tous les scénarios, les

mettent à disposition au format XML de façon à ce que l'utilisateur puisse le cas échéant faire les opérations nécessaires. Cela correspond à un mode "expert".
Le but final étant à terme très clairement d'abstraire cette logique manuelle à travers une forme d'intelligence de code.