



Number-Phone-FR

détails d'implémentation

Journées Perl 2011

2011-06-25

Olivier Mengué
Paris.pm
dolmen@cpan.org
<http://o.mengue.free.fr/>

Number::Phone::FR

- Validation de numéros de téléphone en France
- Intégré au framework Number::Phone
 - Même interface
 - Number::Phone->new()
 - Chargé automatiquement par Number::Phone
- Données de l'ARCEP

Number::Phone

- is_valid
- is_allocated
- is_geographic
- is_fixes_line
- is_mobile
- is_pager
- is_iphone
- is_tollfree
- is_specialrate
- is_adult
- is_corporate
- ...
- subscriber
- operator
- format
- ...

Number::Phone::FR::Full

- Données de l'ARCEP : un fichier Excel des blocs de numéros de téléphone
 - 15172 lignes
 - 2,7 Mo
 - Colonnes utiles : préfixe du bloc, et nom de l'opérateur
- Générer des regexps

Number::FR::Full::RE_FULL

```
(?: (?-xism: (? (?: 3651) ? (? [047-9] | 16 (? : 1 [0-3578] | 2 [1-357] | 3 [03578] | 5 [04679] | 6 [014-6] | 4 [04] | 9 [09] | 03 | 70) ) | (? : 00 | \+ ) 33) ) (?-xism: (? (?: 5 (? : 9 (? : 0 (? : 4 [0-24-8] | 5 [0-26-9] | [289] [0-9] | 0 [2-5] | 1 [0-3] | 6 [0189] | 3 [28] | 7 [07] ) | 4 (? : 3 [0-57-9] | 2 [0-2457-9] | 4 [3-9] | 6 [0-3] | 9 [0147] | 5 [7-9] | 10) | 6 (? : 4 [024-8] | 3 [05-9] | [56] [0-9] | 9 [4-8] | 0 [2-5] | [12] 0 [7 [0-9] | 89) ) | 0 (? : 0 (? : 6 [24-9] | [57] [7-9] | 3 [0-3] | 1 [015] | 4 [04] | [89] 0 [0 [0-9] ) | 1 (? : 6 [0-26] | 5 [0-9] | 23 | 44 | 76) | (? : 9 [0-3] | 6 [0-9] ) [0-9] | 7 (? : [0-6] [0-9] | 80) | 5 (? : 9 [0-2469] | 55 | 80) | 4 (? : 0 [135] | 41) | 8 (? : 5 [05] | 41) | 3 (? : 05 | 66) | 244) | 6 (? : 4 (? : 1 [0-8] | 4 [01] | 0 [0-9] | 33 | 64 | 76 | 88) ) | [1-35] [0-9] [0-9] | 7 [0-9] [0-9] | 8 (? : 7 (? : [0-24-9] [0-9] | 3 [0-24-9] ) | 6 (? : 0 [0-5] | 10 | 65) | [12] [0-9] [0-9] ) | 3 (? : 2 (? : 0 [0-46-9] | 3 [12] | 8 [08] | 1 [0-9] | 20) | 4 [0-9] [0-9] | [135] [0-9] [0-9] ) | 1 (? : 9 (? : [0-6] [0-9] | 7 [0-6] | 8 [08] | 99) | [67] [0-9] [0-9] ) | 4 (? : [24-69] [0-9] | [07] [0-9] [0-9] ) | 5 [35-9] [0-9] [0-9] | (? : 24 | 79) [0-9] [0-9] ) | 4 (? : 0 (? : 0 (? : 6 [3-9] | 1 [0-5] | 7 [05-9] | [04] [0-9] | [389] 0 | 59) | 1 (? : 6 [0-6] | 5 [0-9] | 00 | 23 | 44 | 76) | 4 (? : 0 [03-5] | 4 [1-3] | [89] 0 | 66) | 8 (? : [0-7] [0-9] | 8 [0-2] ) | 3 (? : 1 [36] | 44 | 66) | 5 (? : 7 [78] | 44 | 66) | 2 (? : 6 [5-7] | 44) | [67] [0-9] [0-9] | 9 (? : 40 | 88 | 99) ) | 8 (? : 4 (? : 3 [0-5] | 8 [024-68] | [0-26] [0-9] | 4 [78] | 50 | 76 | 98) | 5 (? : 2 [0-3] | [01] [0-9] | 8 [08] | [46] 4 | 38) | 9 (? : [0-24-9] [0-9] | 3 [014-9] ) | 2 (? : [0-8] [0-9] | 9 [0-289] ) | [01368] [0-9] [0-9] | 799) | 2 (? : 2 (? : 1 [0-6] | 4 [2-4] | 2 [23] | 0 [0-9] | 32 | 84) | 0 (? : [1-4] 0 | 9 [45] | 0 [0-9] | 55 | 61 | 88) | 8 (? : 0 [0-7] | 10) | [67] [0-9] [0-9] ) | 1 (? : 1 (? : [0-8] [0-9] | 9 [0-38] ) | 5 (? : [0-3] [0-9] | 50) | 3 [0-9] [0-9] ) | 4 (? : 4 (? : 0 [0-2] | 44 | 88) | 2 [0-9] [0-9] | 3 [0-9] [0-9] ) | 3 (? : [278] [0-9] [0-9] | [04] [0-9] [0-9] ) | 5 (? : 0 [0-9] [0-9] | [67] [0-9] [0-9] | 800) | 6 (? : [6-8] [0-9] [0-9] | [39] [0-9] [0-9] ) | [79] [0-9] [0-9] [0-9] ) | 9 (? : 7 (? : 4 (? : 0 [1-9] | 8 [0-689] | 4 [0-6] | 7 [135-9] | 5 [0-5] | 1 [5-9] | [236] [0-9] | 91) | 6 (? : [06] [1-9] | 1 [0-357] | 7 [0-367] | 3 [0-37] | 5 [0-2] | 2 [27] | 4 [17] | 9 [0-9] ) | 1 (? : 5 [0-35-9] | [2-4689] [0-9] | 1 [0-5] | 7 [013] | 0 [01] ) | 7 (? : 5 [0-46-9] | 8 [0-79] | [0134679] [0-9] | 2 [0-2] ) | 0 (? : [0-489] [0-9] | 6 [0-3] | 7 [0579] | 5 [014] | 8 (? : 0 [0-5] | 2 [0-3] | 4 [0-2] | 7 [0-9] | 35 | 97) | 3 (? : 8 [24-7] | 3 [0-24] | 0 [0-2] | [45] [0-9] | 73) | 5 (? : [02-9] [0-9] | 1 [0-24-9] ) | 9 (? : [0-8] [0-9] | 9 [0-24-79] ) | 2 (? : [1-69] [0-9] | 0 [9] | 0 [9] | 72) ) | 8 (? : 0 (? : [12] [0-9] | [4-8] 0 | 0 [018] | 3 [05] | 98) | 8 (? : 2 [0-2] | 9 [089] | 8 [0-9] | 11 | 66 | 77) | 2 (? : [2-6] [0-9] | 1 [01] | 82 | 98) | 7 (? : 65 | 87 | 98) | 1 [0-9] [0-9] | 6 (? : 86 | 98) | 31 [1-4] | [45] 98 | 9 [0-9] [0-9] ) | 00 (? : [01] [0-9] | 20 | 65) | [56] [0-9] [0-9] [0-9] ) | 6 (? : 9 (? : 3 (? : 4 [0-7] | 9 [0-47] | 0 [0-4] | [13] [03] | [25-8] 0) | 0 (? : 8 [0-8] | [02] [0-5] | 7 [1-7] | [13-69] [0-9] ) | 4 (? : [04] [0-7] | 2 [0-46-9] | 1 [0-5] | 9 [0-9] | 38) | 6 (? : [0-47-9] [0-9] | [56] [01] ) | [589] [0-9] [0-9] | 2 [0-9] [0-9] ) | 3 (? : 9 (? : 0 [03-79] | 6 [015-9] | 1 [019] | 2 [0-9] | 39 | 40 | 94) | 6 (? : [5-9] [0-9] | [0-4] [0-9] ) | 8 (? : [1-9] [0-9] | 00) | [0-57] [0-9] [0-9] ) | 4 (? : 1 (? : 6 [0-24-9] | [79] [0-9] ) | 9 (? : [0-46-8] [0-9] | [59] [0-9] ) | (? : [235-8] [0-9] | 4 [0-9] ) [0-9] | 0 (? : [1-9] [0-9] | 0 [0-9] ) | 0 (? : (? : [137-9] [0-9] | 2 [0-5] | [45] [0-9] ) [0-9] | 0 (? : [0-36-9] [0-9] | 4 [1-5] | 5 [124] ) | 6 (? : [0-4] [0-9] | [5-9] [0-9] ) | 5 (? : (? : [0-2457-9] [0-9] | 3 [0-9] | 0-9) | 6 (? : [0-59] [0-9] | 6 [6-9] ) ) | [126-8] [0-9] [0-9] [0-9] ) | 2 (? : 0 (? : 0 [0-35-8] | 3 [034] | 7 [089] | 4 [04] | 5 [7] | 66 | 80) | 1 (? : 6 [0-36] | 5 [0-9] | 23 | 44 | 76) | 4 (? : 0 [1-4] | 1 [12] | 44 | 56) | 2 (? : 6 [236] | [89] 0 | 44) | 8 (? : [01] [0-9] | 80) | [67] [0-9] [0-9] | 9 (? : [67] 1 | 55) | 344) | 5 (? : 2 (? : 1 [0-8] | 9 [0-7] | 8 [4689] | 4 [248] | 7 [236] | 0 [0-9] | 62) | 7 (? : 4 [0-68] | 2 [0-25689] | [013] [0-9] | 7 [56] | 8 [48] ) | [14] [0-9] [0-9] | [036] [0-9] [0-9] ) | 6 (? : 2 (? : 7 [0-57] | [2-59] [0-9] | 1 [0679] | 6 [26-8] | 0 [126] | 8 [089] ) | 1 (? : 6 [0-6] | [0-5] [0-9] | 88) | 9 (? : 6 [0-4] | 50) ) | 4 (? : 6 (? : 0 [0-6] [0-9] | 8 [0248] | 7 [06] ) | [01378] [0-9] [0-9] | [459] [0-9] [0-9] ) | 7 (? : 8 (? : 0 [0-8] [0-9] | 9 [0-38] ) | [267] [0-9] [0-9] ) | 3 (? : [1-3578] [0-9] [0-9] ) | [046] [0-9] [0-9] ) | 2 (? : [37-9] [0-9] [0-9] | 2 [0-9] [0-9] ) | 9 (? : [6-9] [0-9] [0-9] | 0 [0-9] [0-9] ) | (? : 1 [48] | 85) [0-9] [0-9] ) | 1 (? : 0 (? : 1 (? : 2 [1-489] | 6 [0-46] | 9 [02-4] | 7 [67] | [48] 0 | 5 [0-9] ) | 4 (? : 0 [124-8] | 4 [0-3] | 8 [1-4] | 7 [12] | 66) | 0 (? : 1 [0-57-9] | [02-8] [0-9] | 9 [0-3] ) | 8 (? : [0-3] [0-9] | 8 [01] | 66 | 92) | 3 (? : 1 [367] | 0 [03] | 33 | 66) | 2 (? : [02] [01] | 6 [06] | 80) | 9 (? : 8 [18] | 9 [19] | 66 | 77) | [5-7] [0-9] [0-9] ) | 8 (? : 0 (? : [23] [1-9] | [014-9] [0-9] ) | 4 (? : [25] [0-4] | [014] [0-9] | 8 [48] | 61) | 1 (? : [0-57-9] [0-9] | 6 [02-9] ) | [23] [0-9] [0-9] ) | 4 (? : 8 (? : [02-57-9] [0-9] | [16] [0-9] ) | 9 (? : [0-689] [0-9] | 7 [0-9] ) | [0-7] [0-9] [0-9] ) | (? : (? : 5 [035-8] | 6 [0149] | 3 [049] | 9 [01] ) [0-9] | 24 [12] ) [0-9] | 7 (? : 1 (? : [1-8] [0-9] | [09] [0-9] ) | [02-9] [0-9] [0-9] ) | 3 (? : 0 (? : 1 (? : 6 [0-367] | 5 [0-9] | 00 | 23 | 44 | 76) | 0 (? : 4 [04] | [08] 0 | 33 | 54 | 66) | 3 (? : 6 [78] | [89] 0 | 16 | 44) | 8 (? : [01] [0-9] | 8 [01] ) | 4 (? : 0 [1-4] | 4 [12] ) | [67] [0-9] [0-9] | 2 (? : 00 | 44) | 9 (? : 88 | 99) ) | 7 (? : 2 (? : [0-3] [0-9] | 8 [238] | 4 [01] | 5 [07] | 7 [23] | 63) | 3 (? : 4 [0-4] | [0-3] [0-9] | 8 [089] | 7 [23] | 9 [89] ) | [01] [0-9] [0-9] | 40 [01] ) | 6 (? : 7 (? : 1 [0-79] | 2 [01] | 8 [48] | 0 [0-9] | 67 | 76) | 5 (? : 6 [0-8] | [0-5] [0-9] | 7 [0-3] ) | [0-4689] [0-9] [0-9] ) | 5 (? : 2 (? : 5 [0-8] | [0-4] [0-9] | 8 [048] | 6 [01] ) | [1457-9] [0-9] [0-9] ) | (? : [28] [0-9] | 9 [01] ) [0-9] [0-9] | 39 (? : 8 [028] | [01] [0-9] | [29] 0) | 4 (? : 4 [0-9] [0-9] | 5 [0-9] [0-9] ) | 10 [0-9] [0-9] ) | 7 (? : 5 (? : 7 (? : [0187-9] | [56] [0-9] | 17 [78] ) | 5 (? : 6 [6-9] | 5 [0-9] ) | 0 (? : [056] [0-9] | 70) | 1 [0-5] [0-9] | 20 [01] | 432) | 8 (? : (? : 8 [0-5] | 7 [0-9] | 100) [0-9] | 6 (? : 1 [1-9] [0-9] | 0 [0-9] ) | 6 (? : 0 (? : 1 [1-9] [0-9] | 0 [0-9] ) | [12] [0-9] [0-9] ) | 7 (? : [78] [0-9] | 0 [0-9] ) | 0-9) | 8 (? : (? : 9 (? : 9 (? : [02] [0-689] | 1 [0-79] | 6 [0235689] | 7 [015-9] | 9 [02-47-9] | 3 [023689] | 5 [04-79] | 8 [017-9] | 4 [0469] ) | 1 (? : 3 [0235-9] | 5 [014-79] | 10 [0-47] | 2 [02-6] | 6 [025-79] | 1 [12468] | 4 [3-68] | 7 [017-9] | 8 [049] | 9 [279] ) | 0 (? : 0 [0-48] | 5 [04-8] | 6 [0-24-6] | 8 [014-69] | 2 [02-5] | 7 [01578] | 9 [0-257] | 3 [2369] | 4 [246] | 12) | 7 (? : 7 [0-24-9] | 1 [023689] | 2 [1-49] | 5 [4-79] | 0 [0134] | 3 [27-9] | 6 [2569] | 9 [0379] | 4 [469] | 89) | 3 (? : 7 [0137-9] | 3 [0-269] | 0 [1246] | 5 [5-79] | 9 [0359] | 1 [168] | 2 [346] | 6 [25] | 8 [47] | 46) | 8 (? : 0 [0-248] | 7 [0-258] | 6 [0256] | 8 [0369] | 2 [034] | 5 [679] | 1 [69] | 3 [29] | 4 [06] | 98) | 2 (? : 1 [0-689] | 7 [0-24-9] | 5 [024-9] | [02-469] [0-9] | 8 [0-2489] ) | 2 (? : 1 (? : [14] [013-9] | 0 [0-57-9] | 8 [02-9] | 3 [01346-8] | 5 [0-57] | 2 [0-35] | 6 [1256] | 7 [0457] | 9 [0-9] ) | 5 (? : 6 [0-35-9] | 2 [0-24-689] | 5 [0-69] | 1 [0-6] | 4 [0457-9] | 7 [0-2457] | 9 [024579] | [038] [0-9] ) | 6 (? : 0 [0-689] | 5 [0-35-79] | 4 [0-6] | 8 [0-24-68] | 2 [0246-8] | 1 [01346] | 9 [05-7] | [367] [0-9] ) | 0 (? : 6 [0-69] | [79] [02457] | 4 [0-58] | 5 [03-8] | [0-38] [0-9] ) ) | 4 (? : 2 (? : 1 [0-46-9] | 2 [0-79] | 5 [457-9] | 3 [346] | 6 [68] | 7 [06] | 9 [07] | 0 [0-9] | 44) | 0 (? : 3 [013-9] | 4 [0-4] | [0-2] [0-9] | 5 [045] | 6 [046] | 7 [06] | 8 [16] | 90) | 1 (? : 0 [013-5] | 4 [12] | 11) | 3 (? : 0 [0-389] | 99) ) | 6 (? : 0 (? : 0 [0-24-79] | 1 [0-25-9] | 4 [0-24-7] | 9 [0-469] | 6 [0-2458] | 7 [014589] | 8 [0246-8] | 3 [0378] | 5 [0179] | 2 [268] ) | 8 (? : 3 [02389] | 4 [0-26] | 9 [0-29] | 7 [089] | 8 [468] | 0 [17] | 5 [27] | 6 [08] | 11 | 22) ) | 1 (? : 1 (? : 1 [0-8] | 6 [0-79] | 1 [0-24-9] | 5 [0-35-79] | [02-489] [0-9] ) | 9 (? : 0 [25] | 1 [09] ) | 0 [0-9] [0-9] ) | 5 (? : 0 [0-2469] | 5 [05] | 8 [58] | 28 | 40 | 66) [0-9] | 8428) [0-9] | 0 (? : (? : 5 (? : 2 [013-9] | 4 [0-79] | 9 [0-35-9] | 3 [02-8] | 5 [0-7] | 6 [0235-79] | 7 [0135-7] | 8 [01568] | [01] [0-9] ) | 9 (? : 0 [0-368] | 1 [023] | 3 [268] | [489] 0 | 54) | (? : 3 [038] | 0 [0-9] ) [0-9] | 1 (? : 6 [39] | 86) ) [0-9] | 880 [0135-9] ) | 36 (? : (? : 0 [1-35-9] | 2 [4-9] | 4 [37-9] | 5 [49] | 6 [02] ) [0-9] | 7 (? : [35] [0-9] | 47) ) ) [0-9] [0-9] [0-9] ) | (? : (?-xism: 3 (? : 2 (? : 0 [0-79] | 1 [0-46-9] | 4 [0-689] | 6 [0-24-9] | 7 [0-689] | 9 [0-35-9] | 5 [0-246-9] | 2 [0-9] | 3 [0-9] | 8 [0-9] ) | 9 (? : 3 [02-9] | 9 [0-689] | 1 [02-689] | 2 [0-24-69] | 7 [0-3579] | 4 [014589] | 8 [07-9] | 5 [059] | 6 [069] | 0 [0-9] ) | 6 (? : 4 [0-24-9] | 6 [0-689] | 5 [0-24-8] | 9 [024-69] | 7 [026-8] | 8 [068] | 0 [0-9] | 2 [0-9] | 3 [0-9] ) | 0 (? : 0 [0-369] | 2 [02356] | 1 [0258] | 8 [0679] | 3 [036] | [467] 0 | 5 [58] | 19 [09] ) | 1 (? : 1 [02-47] | 0 [0136] | 3 [0-3] | 2 [023] | 50 | 88) | (?-xism: 1 (? : 1 (? : 8 (? : 7 (? : 1 [0-3] | 00 | 77) | (? : 31 | 88) 8 | 2 (? : 18 | 22) | 00 [06-8] | 612) | [259] | 6000) | 0 (? : 2 [02-8] | 0 [0-24579] | 6 [148] | 9 [079] | 3 [34] | 4 [34] | 7 [07] | 8 [08] | 1 [0-9] | 55) | [578] ) ) ) )
```

Deux implémentations fournies

- **Simple** :
 - Vérifie juste les formats : nombre de chiffres selon le préfixe
 - Écrite « à la main »
 - Code vérifiable par un humain
- **Full** (« complète ») :
 - Hérite
 - Validité des préfixes, selon données de l'ARCEP
 - Donne l'opérateur contrôlant le bloc
 - Grosses regexps : plus de code chargé, plus lent
 - Code généré
 - Faire confiance à la suite de tests

Problématiques

- **Number : : Phone** ne permet qu'une implémentation par zone téléphonique
 - Comment permettre à l'utilisateur de choisir l'implémentation ?
 - Sélectionner une implémentation par défaut
- Maintenir les données à jour par rapport à l'ARCEP
 - Comment publier facilement un module à jour ?
 - Comment permettre à quelqu'un d'autre de faire cette maintenance ?
- Numéros de versions
 - Code / Données

Sélection de l'implémentation

- Deux classes pour les implémentations :
 - `Number::Phone::FR::Simple`
 - `Number::Phone::FR::Full`
- Une classe enveloppe : `Number::Phone::FR`
- Sélection de la classe à utiliser dans le constructeur
- Plusieurs possibilités :
 - Une variable globale *local*-isable : `$Number::Phone::FR`
 - Une seule implémentation pour tout le programme
 - Contraignant pour les applications à multiple composants indépendants (ex : serveur web)
 - Une sélection par package, à l'import
 - Solution retenue

Sélection de l'implémentation

```
package MonPackage1;
use Number::Phone::FR 'Simple';

say ref Number::Phone->new('+33148901515');
# Number::Phone::FR::Simple

package MonPackage2;
use Number::Phone::FR 'Full';

say ref Number::Phone->new('+33148901515');
# Number::Phone::FR::Full
```

Sélection de l'implémentation

Par la ligne de commande

```
perl -MNumber::Phone::FR=Full -E "say ref  
Number::Phone->new( '+33148901515' )"
```

Par l'environnement

```
export PERL5OPT=-MNumber::Phone::FR=Full
```

L'astuce

- Une fonction « import »
 - Charge l'implémentation désirée
 - Enregistre la sélection pour le package appelant : `caller`
 - `$pkg2impl{ (scalar caller) } = $_[1]`
- Dans le constructeur recherche dans la pile d'appels (« `caller` ») jusqu'à trouver un package avec une sélection mémorisée

Génération du code `Number::Phone::Full`

- Récupération des nouvelles données de l'ARCEP
 - `LWP::UserAgent`
- Lecture des données :
 - `Spreadsheet::ParseExcel`
- Construction des regexps :
 - `Regexp::Assemble`
 - `Regexp::Assemble::Compressed` (sous-classe)
 - Patch des regexps obtenues : `\d` → `[0-9]`
- Génération du source
 - `Template::Toolkit`

Production d'une distribution

```
perl Build.PL
./Build update
# L'update a provoqué un changement de version
perl Build.PL
./Build dist
./Build disttest

# Upload sur le CPAN
```

Extension de `Module::Build`

- La commande `update` est une extension de `Module::Build` embarquée dans la distribution :
 - `inc/Build/Author.pm`
 - `sub ACTION_update`
- L'extension n'est chargée qu'en mode « auteur »
 - Existence d'un répertoire `.git`
 - Les dépendances de l'extension n'impactent pas les simples utilisateurs de la distribution (`./Build install`)

Numéros de version

- **0.0211172**
- **m.nnaajjj**
- **m** : majeur du code
- **nn** : mineur du code
- **aa** : année des données
- **jjj** : jour des données, dans l'année
 - **Last-Modified** du fichier ARCEP

TODO

- Plus de tests
 - Seulement 827 tests

file	stmt	bran	cond	sub	time	total
<u>blib/lib/Number/Phone/FR.pm</u>	82.1	54.5	50.0	59.5	0.0	71.4
<u>blib/lib/Number/Phone/FR/Full.pm</u>	95.2	n/a	n/a	88.9	0.0	93.3

- Doc en français : POD2::FR::Number::Phone::FR
- Optimisation des regexps : des idées, des patches ?
- Dist::Zilla ?
 - Est-il possible d'embarquer des plugins dans la distrib ?

-Ofun

Questions ?



Olivier Mengué

dolmen@cpan.org
<http://o.mengue.free.fr/>