Théorie de la technique des pistes en sudoku

Par Robert Mauriès (*)

Introduction

L'objet de ce document est de présenter et de démontrer les propriétés essentielles de la théorie globale que j'ai développée et appelée "technique des pistes", permettant de résoudre les grilles de sudoku 9x9 classiques considérées dans leur sens le plus large, à savoir qu'on ne se limite pas à des grilles à solution unique.
La suite de ce document suppose acquises de la part du lecteur les règles, les définitions et les techniques simples de résolution des grilles de sudoku. Il suppose aussi une connaissance minimum du langage mathématique et de la théorie des ensembles.

1. Terminologie et conventions

Précisons le sens donné à certains termes utilisés dans ce document.

- Une ***grille*** (sous-entendu grille sudoku) est composée de 9x9=81 cases disposées dans 9 lignes (Li), 9 colonnes (Cj) et 9 blocs (Bk) de 9 cases dans certaines desquelles sont disposés K nombres (K<81) de valeur 1 à 9, appelés des ***dévoilés***.

- Une ***zone*** (sous-entendu zone sudoku) est une ligne, une colonne ou un bloc de la grille sudoku.

- Une ***occurence*** est un nombre de valeur 1 à 9.

- **R** désigne la règle de jeu, qui stipule que l'on doit disposer, dans les cases non occupées par les dévoilés, une occurrence et une seule par case et par zone.

- Un ***candidat*** (sous-entendu candidat potentiel) noté $\mathbf{A_i}$ est une des occurences pouvant être disposés, en appliquant R, dans une case de la grille sudoku qui n'est pas encore résolue. Deux candidats sont différents dès lors qu'ils ne sont pas dans la même case, ou s'ils sont dans la même case n'ont pas la même occurence.
On désigne par $\mathbf{G}=\{A_i , i=1….., N\}$ l'ensemble de ces candidats pouvant être disposés dans les différentes cases libres de la grille, avec $N<(81-K)x9$.

Un candidat est repéré par sa valeur et sa position (ligne-colonne) dans la grille. Par exemple 5L2C6 désigne le candidat de valeur 5 dans la case située à l'intersection de la ligne 2 et de la colonne 6.

- ***Placer*** un candidat sur une grille est l'opération consistant à disposer ce candidat seul dans une case ou d'une zone de la grille.

- ***Eliminer*** un candidat sur une grille est l'opération consistant à supprimer ce candidat d'une case de la grille.

- Une ***solution*** d'une grille sudoku, si elle existe, est l'ensemble $\mathbf{S_n} = \{A_i \in G,\ i=i_1, i_2, \ldots, i_p,\ p=81-K\}$ des candidats placés dans toutes les cases de la grille de telle manière que la règle R soit respectée. Une grille peut avoir une solution (**n**=1), plusieurs solutions (**n** = 1, 2, ...) ou aucune solution.

- Un ***candidat solution*** est un candidat appartenant à une S_n**.**
A contrario, un ***candidat non solution*** est un candidat n'appartenant à aucune S_n**.**

- Une **technique de résolution** est une démarche (raisonnement) logique permettant de placer ou d'éliminer des candidats sur une grille en appliquant R. On désigne par **TR** l'ensemble des techniques de résolution permettant de résoudre une grille, par **TE** (techniques élémentaires) une partie de ces techniques incluant les plus simples que sont : la recherche des candidats uniques par balayage des zones, la recherche des alignements et la recherche des ensembles fermés (doublet, triplet, quadruplet, etc). L'ensemble de ces trois techniques constituant les **techniques de base** désignées par **TB.**

- On désigne par $\mathbf{E} = \{A_k \in G,\ k=k_1, \ldots, k_m\}$, où m < N, un ensemble de candidats qui forme un sous-ensemble de l'ensemble G.

- Une ***entité*** de G est l'ensemble formé de tous les candidats d'une même case, ou de tous les candidats de même occurrence d'une même zone

- Deux ensembles disjoints **E1** et **E2** forment ***une paire d'ensembles*** de G lorsque leur réunion est une entité de G.

- Deux candidats forment une ***paire***, lorsque leur réunion est une entité de **G**.

- On dit qu'un candidat A_i ***voit*** un candidat différent A_r lorsque les deux candidats A_i et A_r sont dans la même entité.

2. Piste et antipiste.

Définitions 2-1:

1) Une piste $P(A_k)$ issue d'un candidat A_k est l'ensemble des candidats $A_i \in G$ que l'on placerait avec les TR si A_k était placé.

2) Une antipiste $P'(A_k)$ issue d'un candidat A_k est l'ensemble des candidats $A_i \in G$ que l'on placerait avec les TR si A_k était éliminé de la grille.
On dit que $P'(A_k)$ est l'antipiste de la piste $P(A_k)$.

3) Une piste $P(E)$ issue d'un ensemble de candidats E est l'ensemble des candidats $A_i \in G$ communs à toutes les pistes issues de tous les candidats $A_k \in E$.
On a donc $P(E) = \bigcap_E P(A_k)$, soit $P(E) \subseteq P(A_k)$ $\forall k$.

4) Une antipiste $P'(E)$ issue d'un ensemble de candidats E est l'ensemble des candidats $A_i \in G$ que l'on placerait avec les TR si on éliminait tous les candidats $A_k \in E$.
On dit que $P'(E)$ est l'antipiste de la piste $P(E)$.

Convention :
P et P' sont l'une et l'autre des pistes au sens large, seul leur mode de construction diffère. Aussi et sauf précision expresse, sauf si la distinction des deux est nécessaire, on utilisera dans la suite le terme de "**piste P**" aussi bien pour une piste que pour une antipiste lorsque un résultat énoncé est valable pour l'une ou l'autre de ces deux catégories de pistes.

Remarques importantes:

- Au plan pratique, une piste P est représentée par un marquage graphique sur la grille (couleur ou symbole) comme sur la figure 1 où seules les TB ont été utilisées. Les TB ou les TE seules étant parfois insuffisantes pour identifier tous les candidats possibles d'une piste P, la construction de celle-ci est bloquée. Ce n'est pas pour autant que la piste P se limite à ces quelques candidats, elle est composée aussi des candidats non encore identifiés qui seront trouvés si on utilise d'autres techniques de TR.
Une piste P est donc formée des candidats identifiés et non identifiés par les seules TB ou les seules TE.
On en vient donc à donner la définition suivante.

Définition 2-2:
La trace d'une piste P est la piste $\underline{P}$ construite selon les définitions 2-1 où seules les TB sont utilisées à la place des TR.

- La construction d'une piste, c'est à dire l'identification des candidats qui la composent, pouvant être abordée de plusieurs manières différentes, il n'est pas exclu qu'une piste se compose finalement de plusieurs candidats d'une même case, ou/et de plusieurs candidats de même valeur d'une même zone.

Par exemple, sur la grille partielle ci-dessous, les candidats de la trace $\underline{P}$(1L3C5) sont marqués d'un carré bleu. Selon qu'on construit la piste en suivant les flèches vers la gauche ou vers la droite, on déduit que 1L1C1 et 1L1C9 sont des candidats de P.

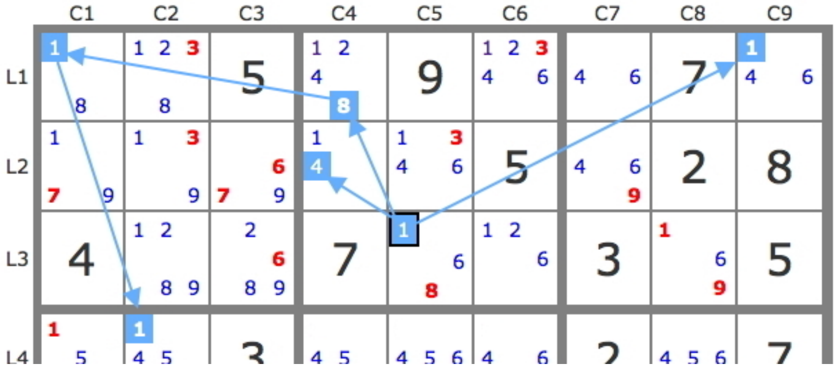

Fig 1 : Tracé partiel de la piste P(1L3C5)

Dans ce cas il est clair qu'une telle piste ne peut pas représenter une solution de la grille puisqu'au final elle est incompatible avec R**,** mais il s'agit bien d'une piste correctement construite.

Enonçons quelques propriétés des pistes utiles pour la suite.

Propriété 2-1 :
Si E est la réunion de plusieurs ensembles E_r, r=1,2, …p, on a :
$P(E) = \bigcap_{E} P(E_r)$ donc $P(E) \subseteq P(E_r)$ $\forall r$,
Les pistes $P(E_r)$ sont appellées les branches de P(E).

En effet, on peut écrire:
$P(E) = \bigcap_{E} P(A_k) = \bigcap_{E} (\bigcap_{E_r} P(A_k)) = \bigcap_{E} P(E_r)$.

Propriété 2-2 :
Si E1 $\subseteq$ E2, alors P(E2) $\subseteq$ P(E1).

En effet, E'1 désignant le complémentaire de E1 dans E2, on a d'après la propriété 2-1 : P(E2)=P(E1)$\cap$P(E'1) $\subseteq$ P(E1).

Propriété 2-3 :
Pour toute piste P, $\forall A_i \in P$ on a : $P(A_i) \subseteq P$.

En effet, si $A_i \in P$ c'est que A_i a été placé en utilisant les TR. En plaçant un candidat A_k par les TR comme un candidat de $P(A_i)$ on ne fait que poursuivre, par définition, la construction de P. Donc $A_k \in P$ aussi. Tout candidat de $P(A_i)$ est donc un candidat de P, soit $P(A_i) \subseteq P$.

Le théorème qui suit exprime la réciprocité qui existe entre piste et antipiste issues des deux candidats d'une paire de candidats ou des deux ensembles de candidats d'une paire d'ensembles.

Théorème 2-1 :
Si E1 et E2 forment une paire d'ensembles, la piste P(E1) issue de E1 et identique à l'antipiste P'(E2) issue de E2, et réciproquement.
En particulier, si deux candidats forment une paire, la piste issue de l'un et identique à l'antipiste issue de l'autre.

En effet :
- l'antipiste P'(E2) est construite en supposant que les candidats de E2 sont éliminés, donc que seuls les candidats de E1 subsistent puisque E1 et E2 forment une paire d'ensembles.
Comme un des candidats A_k de E1 doit être placé pour appliquer R, on peut dire que un des A_k de E1 est un candidat de P'(E2), donc que pour ce A_k on a $P(A_k) \subseteq P'(E2)$, donc que $P(E1) \subseteq P(A_k) \subseteq P'(E2)$.
- Inversement, A_1, A_2, … A_p, étant les candidats formant E1, si E'_k désigne le complémentaire de A_k dans E1$\cup$E2, on peut écrire :
$P'(E'_k) = P(A_k)$ $\forall$ k=1,2..p, puisque seul A_k subsiste dans la case ou la zone contenant la paire d'ensembles.
Or, si un candidat A_i est placé par les TR en supprimant tous les candidats de E2, à fortiori A_i est placé par les TR si on supprime tous les candidats de E'_k, donc (cf définition d'une antipiste) si $A_i \in P'(E2)$ alors $A_i \in P'(E'_k)$, c'est-à-dire que $P'(E2) \subseteq P'(E'_k)= P(A_k)$ et ce $\forall k$,
d'où $P'(E2) \subseteq \bigcap_k P(A_k)=P(E1)$.
En rassemblant les deux inclusions inverses on a : P(E1)= P'(E2), et en inversant les rôles de E1 et E2, P(E2)= P'(E1).

Le cas d’une paire de deux candidats étant un cas particulier où E1 et E2 sont réduits à un seul candidat est établit aussi.

3. Validité et Invalidité d’une piste ou d'une antipiste

Définitions 3-1 :
1) Une piste P est valide, si et seulement si, P n'est composée que de candidats solutions.
2) Une piste P est invalide lorsque le placement des candidats qui la composent conduit à une incompatibilité avec R.

De cette définition découle la propriété suivante, presque comme une évidence.

Propriété 3-1 :
Une piste P est forcément valide ou invalide, mais ne peut pas être les deux à la fois, elle a donc l’un ou l’autre des deux statuts.

En effet, si une piste était à la fois valide et invalide, nous pourrions affirmer que tous ses candidats sont solutions (valide) alors que leur placement ne respecte pas R (invalide), ce qui est absurde d’après la définition d'une solution.
D'autre part, si une piste était ni invalide ni valide, nous pourrions affirmer, ce qui serait contradictoire, que :
- n'étant pas invalide, elle aurait un candidat et un seul dans chaque case et dans chaque zone de la grille, ceci en respectant les règles R par construction, elle constituerait donc une solution par définition, c’est-à-dire qu'elle ne serait composée que de candidats solutions.
- n'étant pas valide, elle ne serait pas composée que de candidats solutions.

Attention :
La notion de validité reposant par définition sur celle de candidat solution, cela suppose implicitement que la grille a au moins une solution lorsqu'on évoque la notion de validité**.** Ce qui n'est pas le cas de l'invalidité.
Ainsi dans le cas d’une grille qui n’a pas de solution, aucune piste ne peut être valide, en conséquence toutes les pistes que l’on peut construire sont forcément invalides et réciproquement.

<u>Pour la suite de ce document, nous supposons sauf mention contraire que les grilles ont au moins une solution</u>.

Définitions 3-2:
1) Une piste est S_n-valide, si et seulement si, $P \subseteq S_n$.
2) Une piste est S_n-invalide, si et seulement si, $P \not\subseteq S_n$.

Etablissons quelques propriétés utiles pour les démonstrations du chapitre 4 consacré aux pistes conjuguées.

Propriété 3-2 :
Une piste P est forcément S_n-valide ou S_n-invalide, mais ne peut pas être les deux à la fois, elle a donc l'un ou l'autre des deux statuts.

En effet, cela résulte du fait qu'une piste ne peut pas à la fois être telle que $P \subseteq S_n$ et $P \not\subseteq S_n$.

Propriété 3-3 :
Pour toute piste P on peut énoncer que :
a) P est valide, si et seulement si, $\exists\, S_n$ telle que P est S_n-valide.
b) P est invalide, si et seulement si, $\forall\, S_n$ P est S_n-invalide.

En effet,
a) si une piste P est valide, P n'est composée que de candidats solutions qui par définition et construction de P sont les candidats d'une même solution. Il existe donc une solution S_n telle que $P \subseteq S_n$, c'est-à-dire que P est S_n-valide. Inversement, si $\exists\, S_n$ telle que P est S_n-valide alors $P \subseteq S_n$ et P n'est composée que de candidats solutions, P est donc valide par définition.
b) si une piste P est invalide et si pour une S_n on avait $P \subseteq S_n$, les candidats de P étant des candidats de S_n, cela signifierait que le placement des candidats de S_n conduit à une incompatibilté avec R, ce qui par définition n'est pas vrai pour une solution. Donc, si une piste est invalide on a $P \not\subseteq S_n$ $\forall S_n$, c'est-à-dire que P est S_n-invalide $\forall S_n$. Inversement, si $\forall S_n$ P est S_n-invalide, alors $P \not\subseteq S_n$ $\forall S_n$ et P ne peut pas être valide car, ce serait une contradiction, P valide entrainerait qu'il existe au moins une S_n telle que $P \subseteq S_n$. Donc P est invalide.

Remarques :
- La notion de validité est, comme l'indique la propriété 3-3, attachée à une solution S_n. Une piste valide est S_n-valide. Une grille ayant plusieurs solutions peut donc avoir plusieurs pistes valides différentes.
- Au contraire, la notion d'invalidité n'est attachée à aucune solution, une piste invalide étant S_n-invalide pour toutes les solutions S_n.

Propriété 3-4 :
Soient deux pistes P1 et P2 telles que P1 $\subseteq$ P2, alors :
a) Si P2 est S_n-valide, alors P1 est S_n-valide.
b) Si P1 est S_n-invalide, alors P2 est S_n-invalide.
c) Si P1 est invalide, alors P2 est invalide.
En effet,
a) La piste P2 étant S_n-valide, nous pouvons écrire que P1 $\subseteq$ P2 $\subseteq$ S_n , soit P1 $\subseteq$ S_n, donc la piste P1 est S_n-valide.
b) La seconde assertion découle de la première, car si P2 était S_n-valide, P1 le serait aussi et ce serait contradictoire.
c) La piste P1 étant invalide, elle est S_n-invalide $\forall$ S_n , donc P2 est S_n-invalide $\forall$ S_n aussi, c'est-à-dire est invalide.

Propriété 3-5 :
P(A_k) étant une piste issue d'un candidat A_k :
a) P(A_k) est S_n-valide, si et seulement si, $A_k \in S_n$.
b) P(A_k) est S_n-invalide, si et seulement si, A_k n'appartient pas à S_n.
c) P(A_k) est invalide, si et seulement si, A_k n'appartient à aucune S_n.

En effet,
a) si P(A_k) est S_n-valide alors P(A_k) $\subseteq$ S_n. A_k étant par construction un candidat de P(A_k), on a $A_k \in S_n$. Inversement, si $A_k \in S_n$, tous les candidats de P(A_k) sont des candidats solutions appartenant à S_n par définition et construction de P, donc P(A_k) $\subseteq$ S_n , c'est-à-dire que P(A_k) est S_n-valide.
b) La seconde assertion découle de la première, par équivalence des contraires .
c) si P(A_k) est invalide alors P(A_k) est S_n-invalide $\forall$ S_n et donc A_k n'appartient à aucune S_n. Inversement, si A_k n'appartient à aucune S_n P(A_k) est S_n-invalide $\forall$ S_n , c'est à dire que P(A_k) est invalide.

Propriété 3-6 :
P(E) étant une piste issue d'un ensemble de candidats E :
a) si un $A_k \in E$ est un candidat solution $\in S_n$, alors P(E) est S_n-valide.
b) si P(E) est S_n-invalide, alors $\forall A_k \in E$, A_k n'appartient pas à S_n.
c) si P(E) est invalide, alors $\forall A_k \in E$, A_k n'appartient à aucune S_n.

En effet :
a) Si un candidat $A_k \in$ E est un candidat solution $\in S_n$, la piste P(A_k) est une piste S_n-valide. Comme P(E) $\subseteq$ P(A_k) $\subseteq$ S_n, P(E) est S_n-valide.
b) La seconde assertion découle de la première, car si un A_k appartenait à S_n , P(E) serait S_n-valide ce qui est contradictoire.

c) P(E) étant invalide, P(E) est S_n-invalide $\forall S_n$, donc $\forall A_k \in E$, A_k n'appartient pas à S_n d'après l'assertion précédente et ce $\forall S_n$.

Propriété 3-7:
P'(A_k) étant une antipiste issue d'un candidat A_k :
a) si P'(A_k) est S_n-invalide, alors A_k est un candidat solution $\in S_n$.
b) si P'(A_k) est invalide, alors A_k est un candidat $\in S_n$ $\forall S_n$.
c) si A_k n'appartient pas à S_n , alors P'(A_k) est S_n-valide.

En effet, E étant le complémentaire de A_k dans la case contenant A_k, on a P'(A_k)=P(E) d'après le théorème 2-1.
a) Si P'(A_k) est S_n-invalide, P(E) l'est aussi et aucun des candidats de E n'appartient à S_n. S_n ayant forcément un candidat solution dans chaque case c'est A_k qui est le candidat solution $\in S_n$.
b) Si P'(A_k) est invalide, alors P'(A_k) est S_n-invalide $\forall S_n$ et A_k est un candidat solution appartenant à S_n $\forall S_n$.
c) La troisième assertion découle de la première, car si P'(A_k) est S_n-invalide alors A_k appartient à S_n ce qui est contradictoire.

Propriété 3-8:
P'(E) étant une antipiste issue d'un ensemble de candidats E :
a) si $\forall A_k \in E$, A_k n'appartient pas à S_n , alors P'(E) est S_n-valide.
b) si P'(E) est S_n-invalide, alors un $A_k \in E$ au moins est un candidat solution $\in S_n$.
c) si P'(E) est invalide, alors pour chaque S_n, un $A_k \in E$ au moins est un candidat solution $\in S_n$.

En effet :
a) si $\forall A_k \in E$ A_k n'appartient pas à S_n , par définition tous les candidats de P'(E) sont des candidats de S_n , donc P(E) $\subseteq S_n$, c'est-à-dire que P'(E) est S_n-valide
b) La deuxième assertion découle de la première, car si aucun A_k n'appartenait à S_n P'(E) serait S_n-valide ce qui serait contradictoire.
c) si P'(E) est invalide, P'(E) est S_n-invalide $\forall S_n$ et d'après l'assertion précédente E a au moins un candidat $\in S_n$ et ce pour chaque S_n.

Conséquences pratiques :

Une conséquence pratique des propriétés 3-5 et 3-6 pour la résolution des grilles est que si P(A_k) est invalide on peut éliminer A_k, ou, si P(E) est invalide on peut éliminer de la grille tous les candidats de E.

Une conséquence pratique de la propriété 3-7 est qu'un candidat $A_{\mathbf{k}}$ peut être placé sur la grille comme seule solution de sa case si P'($A_{\mathbf{k}}$) est invalide.

La propriété 3-9 qui suit exprime la dualité existant entre piste et antipiste issues d'un même ensemble E.

Propriété 3-9 : (dualité piste-antipiste):
La piste P(E) et son antipiste P'(E) ne peuvent pas être simultanément invalides :
a) si P(E) est invalide, P'(E) est S_n-valide ∀ S_n, et réciproquement,
b) si P'(E) est invalide, P(E) est S_n-valide ∀ S_n.

En effet cela résulte des propriétés 3-6 et 3-8:
a) Si P(E) est invalide, alors $\forall A_k \in$ E, A_k n'appartient à aucune S_n, donc l'antipiste P'(E) est S_n-valide et ce ∀ S_n.
b) Si P'(E) est invalide, alors pour chaque S_n, un $A_k \in$ E au moins est un candidat solution $\in S_n$, donc P(E) est S_n-valide et ce ∀ S_n.

Ce résultat est fondamental pour la résolution des grilles comme l'est aussi celui de la propriété 3-10 suivante établissant une autre forme de dualité.

Propriété 3-10 : (dualité de deux ensembles):
E étant la réunion de deux ensembles E1 et E2 distincts, si P'(E) est invalide, alors les pistes P(E1) et P(E2) ne peuvent pas être simultanément invalides :
a) si P(E1) est invalide, P(E2) est S_n-valide ∀ S_n, et réciproquement,
b) si P(E2) est invalide, P(E1) est S_n-valide ∀ S_n.

En effet : L'hypothèse P'(E) invalide signifie que pour chaque S_n un candidat A_k de E au moins est un candidat solution appartenant à S_n.
a) Si P(E1) est invalide, alors $\forall A_k \in$ E1 A_k n'appartient à aucune S_n, c'est donc que pour chaque S_n un au moins des candidats de E2 est un candidat solution $\in S_n$, donc P(E2) est S_n-valide et ce ∀ S_n .
b) Si P(E2) est invalide, alors $\forall A_k \in$ E2 A_k n'appartient à aucune S_n, c'est donc que pour chaque S_n un au moins des candidats de E1 est un candidat solution $\in S_n$, donc P(E1) est S_n-valide et ce ∀ S_n .

A noter que si E1 et E2 doivent être distincts, c'est-à-dire ne pas être composés rigoureusement des mêmes candidats, ils ne doivent pas être nécessairement disjoints.

Les pistes qui satisfont l'une ou l'autre des deux propriétés précédentes jouent un rôle déterminant dans la résolution des grilles. Elles font partie d'une famille de pistes qualifiées de pistes conjuguées , faisant l'objet du chapitre suivant.

4. Pistes conjuguées

Définition 4-1 :
Deux pistes P1 et P2 sont conjuguées lorsque la S_n-invalidité supposée de l'une impose la S_n-validité de l'autre et réciproquement, soit :
- si P1 S_n-invalide, alors P2 S_n-valide, et,
- si P2 S_n-invalide, alors P1 S_n-valide.
Deux pistes conjuguées forment un jeu de pistes, noté JP.

Une conséquence de cette définition est que deux pistes conjuguées peuvent être toutes deux valides, circonstance que l'on rencontre nécessairement dans les grilles à solutions multiples, mais aussi dans les grilles à solution unique.
La propriété suivante précise ce qu'il en est de leurs invalidités respectives.

Propriété 4-1:
Deux pistes conjuguées ne peuvent pas être simultanément S_n-invalides, ni simultanément invalides.

En effet,
- si P1 conjuguée de P2 (resp. P2 conjuguée de P1) est S_n-invalide, alors P2 (resp.P1) est par définition S_n-valide et ne peut donc (propriété 3-2) pas être S_n-invalide. P1 et P2 ne peuvent donc pas être simultanément S_n-invalides.
- si P1 et P2 conjuguées étaient simultanément invalides, nous pourrions écrire que P1 $\not\subset S_n$ et P2 $\not\subset S_n$ donc que P1 et P2 seraient simultanément S_n-invalides, ce qui est impossible. P1 et P2 ne peuvent donc pas être simultanément invalides.

Les deux théorèmes suivant (qu'il faut rapprocher des propriétés 3-9 et 3-10), obtenus en combinant les propriétés 3-6 et 3-8, permettent de trouver au plan pratique la plupart des pistes conjuguées.

Théorème 4-1 :
Une piste P(E) et son antipiste P'(E) sont des pistes conjuguées.

En effet,
- si P(E) est S_n-invalide, alors $\forall A_k \in$ E, A_k n'appartient pas à S_n et donc P'(E) est S_n-valide.
- si P'(E) est S_n-invalide, alors un $A_k \in$E au moins est un candidat $\in S_n$ et donc P(E) est S_n-valide.

Théorème 4-2 :
E étant la réunion de deux ensembles E1 et E2 distincts, si P'(E) est invalide, alors les pistes P(E1) et P(E2) sont conjuguées.

En effet : L'hypothèse P'(E) invalide signifie que pour chaque S_n un candidat A_k de E au moins est un candidat solution appartenant à S_n.
- Si P(E1) est S_n-invalide, alors $\forall A_k \in$ E1 A_k n'appartient pas à S_n, c'est donc que un au moins des candidats de E2 est un candidat solution appartenant à S_n, donc P(E2) est S_n-valide.
- Si P(E2) est S_n-invalide, alors $\forall A_k \in$ E2 A_k n'appartient pas à S_n, c'est donc que un au moins des candidats de E1 est un candidat solution appartenantn à S_n, donc P(E1) est S_n-valide.

Les pistes conjuguées jouissent des propriétés d'interactions suivantes qui sont des outils efficaces de résolutions des grilles sudoku.

Propriété 4-2 (dualité):
Si une piste P1 est invalide, tous les candidats d'une piste P2 conjuguée de P1 sont des candidats solutions communs à toutes les solutions de la grille.

En effet, si P1 est invalide alors P1 est S_n-invalide $\forall S_n$. Si P2 est une piste conjuguée de P1 alors P2 est S_n-valide et ce $\forall S_n$, ses candidats sont donc des candidats solutions communs à toutes les solutions.

Pour établir la deuxième propriété d'interaction, donnons la définition suivante.

Définitions 4-2:
On dit qu'un candidat voit une piste P lorsqu'il voit un candidat faisant partie de la piste P.

Propriété 4-3 (élimination) :
Un candidat qui voit deux pistes conjuguées n'est pas, $\forall S_n$, un candidat solution de S_n et peut être éliminé de la grille.

Soient P1 et P2 les deux pistes conjuguées, et $A_{\mathbf{n}}$ un candidat qui voit P1 et P2. $A_{\mathbf{n}}$ voit donc un candidat $A_{\mathbf{i}}$ de P1 et un candidat $A_{\mathbf{r}}$ de P2
Si $A_{\mathbf{n}}$ était un candidat solution $\in S_n$, $A_{\mathbf{i}}$ et $A_{\mathbf{r}}$ ne seraient pas des candidats $\in S_n$ en vertu de R, donc $P(A_{\mathbf{i}})$ et $P(A_{\mathbf{r}})$ ainsi que P1 et P2 (propriété 3-4) seraient toutes deux S_n-invalides, ce qui est impossible (propriété 4-1). Donc $A_{\mathbf{n}}$ n'est pas un candidat solution de S_n et ce $\forall S_n$, il peut donc être éliminé de la grille.

Propriété 4-4 (validation) :
Un candidat commun à deux pistes conjuguées est un candidat solution commun à toutes les solutions de la grille.

En effet, cette propriété est une conséquence de la propriété 4-3, puisque tous les autres candidats de la case contenant un candidat commun à P1 et P2 voient P1 et P2, et en conséqence peuvent être éliminés $\forall S_n$. Ne subsiste donc plus dans cette case comme candidat pouvant être solution que le candidat commun à P1 et P2.

5. Application

L'application de ces propriétés est un outil puissant de résolution des grilles sudoku, car il est toujours possible de construire des pistes conjuguées sur une grille.

La trace d'une piste P se contruit pas à pas en utilisant les TB et les propriétés de validation ou d'éliminations obtenues par interaction avec une autre piste conjuguée.

En pratique, les candidats de la trace d'une piste peuvent être marqués sur la grille par des symboles (. / *) ou des couleurs comme dans l'exemple qui suit.

Sur la grille à solution unique de la figure ci-dessous que l'on ne peut pas résoudre avec les seules TB, on choisit l'ensemble de candidats E={2L7C9, 3L8C1, 9L8C1}.
<u>P'</u>(E) est marquée avec la couleur verte. Elle est construite partiellement en supposant que les 3 candidats de E sont éliminés.
<u>P</u>(E) issue de E comprend les candidats communs de la piste bleue issue du 2L7C9 et de la piste jaune issue de l'ensemble {3L8C1, 9L8C1} qui forme un ensemble fermé (doublet) jaune avec la paire {3L8C7, 9L8C7}. <u>P</u>(E) est donc formée partiellement de tous les candidats jaune entourés d'un carré bleu.

L'interaction de ces trois pistes permet donc d'éliminer tous les candidats barrés d'un trait rouge car ils voient à la fois P(E) et P'(E) qui sont conjuguées.
Mais si on poursuit la construction P'(E) avec les TB on aboutit à son invalidité (affirmation laissée à votre vérification), ce qui par conséquent rend la piste P(E) valide. Tous les candidats de P(E), c'est à dire tous les candidats jaunes entourés d'un trait bleu, sont solutions de la grille qui se termine alors facilement par les TB.

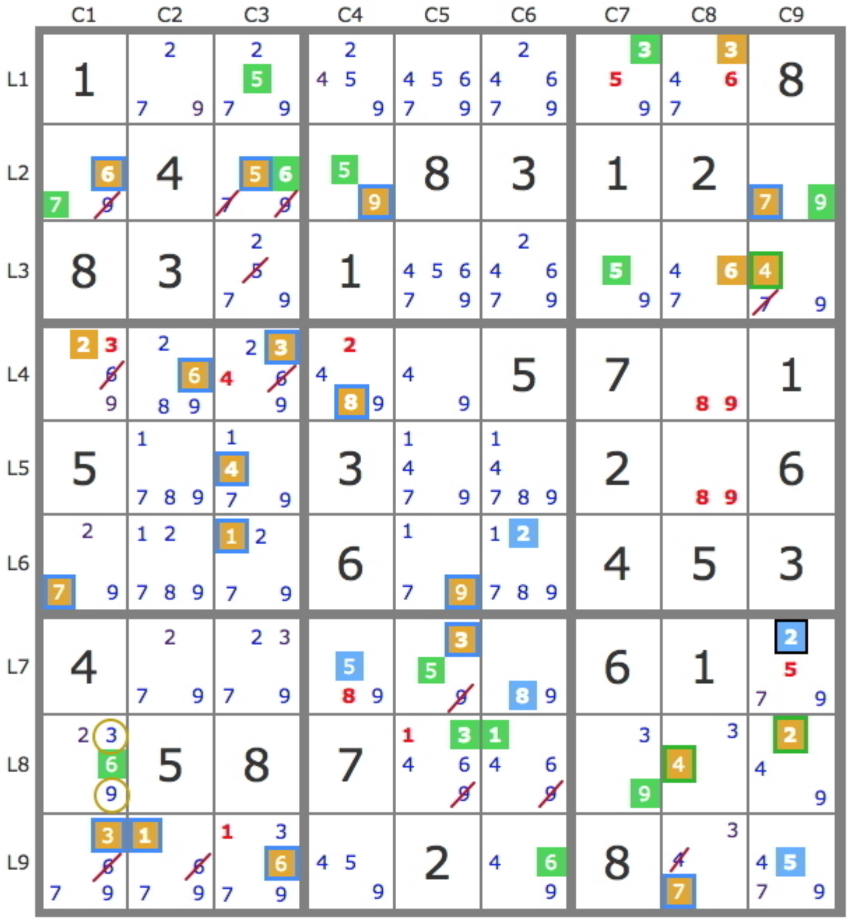

Fig 2 : croisements piste et antipiste

6. Pistes conjuguées et pistes opposées

Avant de définir la notion de pistes opposées, énonçons une propriété relative à la validité des pistes qui sera utile pour les démonstrations des propriétés attachées aux pistes opposées.

Propriété 6-1:
Si faire l'hypothèse P1 est S_n-valide entraîne que P2 est S_n-valide, alors P2 $\subseteq$ P1.

Ce résultat provient du mode de construction d'une piste. Supposer (hypothèse) ou constater qu'une piste est S_n-valide revient à placer les candidats de la piste sur la grille par les TR comme si ils étaient tous des candidats solutions de S_n (voir définition d'une piste). Si de placer ceux de P1 permet de dire aussi que l'on peut placer ceux de P2, alors c'est que P1 se construit aussi avec les candidats de P2. Donc P2 $\subseteq$ P1.

Définition 6-1 :
Deux pistes Q1 et P1 sont opposées lorsqu'un candidat de Q1 voit un candidat de P1 et réciproquement.

A noter qu'en général deux pistes opposées ne sont pas forcément conjuguées car elles peuvent être toutes deux invalides.

Propriété 6-2:
Deux pistes Q1 et P1 opposées ne peuvent pas être toutes les deux S_n-valides.

En effet, soient A_i et A_r deux candidats respectifs de Q1 et P1 qui se voient. Si Q1 et P1 étaient S_n-valide, les candidats de Q1 et de P1 étant tous des candidats de S_n , A_i et A_r seraient des candidats de S_n. S_n serait dans ce cas une solution qui contient deux candidats A_i et A_r qui se voient (même entité), ce qui est impossible, une solution n'ayant qu'un candidat par entité. Q1 et P1 ne peuvent donc pas être S_k-valides toutes les deux.

A noter, que dans une grille à solutions multiples, donc ayant plusieurs solutions, deux pistes peuvent être valides et opposées, mais elles ne seront pas valides pour la même solution. Autant dire que cette propriété doit être utilisée avec précaution sur ce genre de grille.

Théorème 6-1 :
Si Q1 est une piste opposée à une piste P1, toute piste P2 conjuguée de P1 est incluse dans Q1 (P2 $\subseteq$ Q1).

En effet, P1 et P2 étant conjuguées, supposer que Q1 est S_n-valide, entraîne que P1 qui lui est opposée est S_n-invalide, donc que P2 est S_n-valide. Ainsi, faire l'hypothèse "Q1 est S_n-valide " entraînant "P2 est S_n-valide " , ce $\forall S_n$, on a donc P2 $\subseteq$ Q1 d'après la propriété 6-1.

Corollaire 6-1-1 :
Si deux pistes conjuguées Q1 et Q2 sont opposées à une même piste P1, toute piste P2 conjuguée de P1 est S_n-valide $\forall S_n$.

En effet,
Si Q1 et Q2 sont opposées à P1, pour une piste P2 conjuguée de P1 on a d'après le théorème 6-1 : P2 $\subseteq$ Q1 et P2 $\subseteq$ Q2.
Comme, $\forall S_n$, une des deux pistes Q1 ou Q2 est S_n-valide, d'après la propriété 3-4 P2 est S_n-valide $\forall S_n$.

Corollaire 6-1-2 :
P1 et Q1 étant deux pistes opposées, deux pistes P2 et Q2 respectivement conjuguées de P1 et Q1 sont conjuguées.

En effet, d'après le théoreme 6-1, on a P2 $\subseteq$ Q1 et Q2 $\subseteq$ P1.
Si P2 est supposée S_n-invalide cela entraîne que Q1 est S_n-invalide aussi (propriété 3-4), donc Q2 conjuguée de Q1 est S_n-valide. De même, si Q2 est supposée S_n-invalide cela entraîne que P1 est S_n-invalide aussi, donc P2 conjuguée de P1 est S_n-valide. P2 et Q2, satisfaisant la définition des pistes conjuguées, sont donc conjuguées.

7. Application

Voici un exemple d'application de ces notions liées aux pistes conjuguées et aux pistes opposées.

Sur la grille de la figure suivante, on trace partiellement un jeu de pistes conjuguées bleue/jaune issues de la paire 4B2 et un jeu de pistes conjuguées verte/violette issues de la paire 6B3.
Les deux pistes verte (Q1) et violette (Q2) sont opposées à la piste jaune (P1), pour Q1 dans B8 et pour Q2 dans C1.
On peut donc, en vertu du corollaire 6-1-1, valider les 3 candidats de la piste bleue.

La grille peut-être de cette manière entièrement résolue et pour cela nous vous renvoyons à la solution donnée sur cette page web : http://www.assistant-sudoku.com/Grille_Resolue.php?RID=303

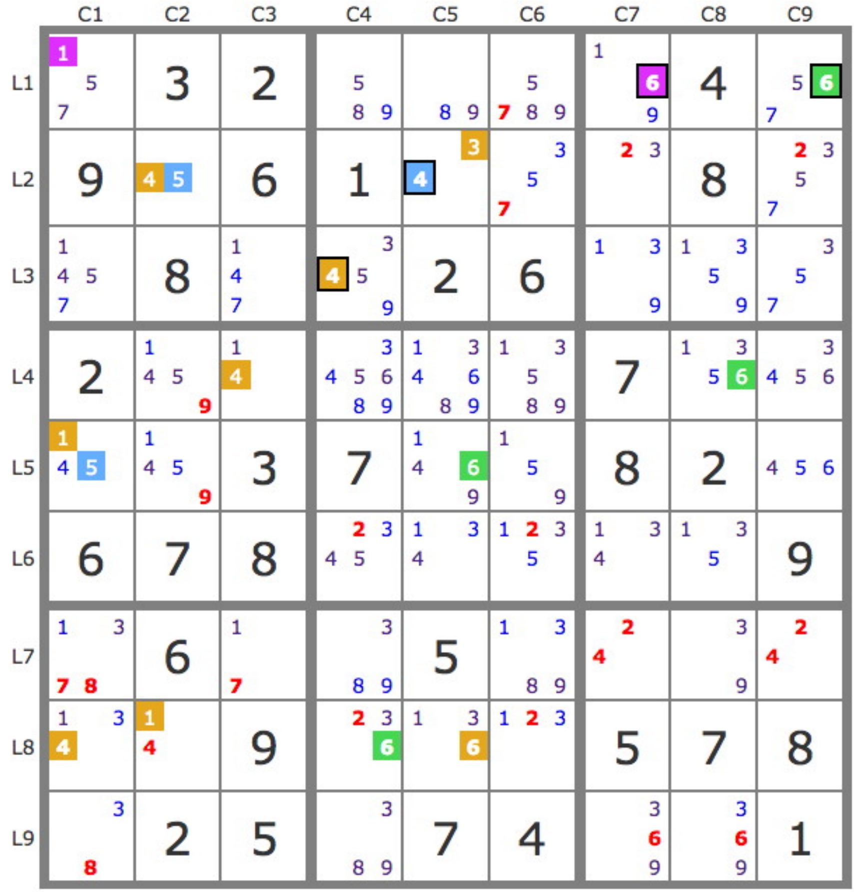

Fig 3 : interactions de pistes opposées et conjuguées

8. Extension d’une piste.

Sur les grilles difficiles, la représentation d’une piste P se limite à sa trace P car sa construction par les TB ou les TE est rapidement bloquée. Pour contourner cette difficulté on utilise une extension de cette piste.
A cet effet, définissons d’abord ce que l’on appelle une P-piste associée à une piste P1.

Définitions 8-1 (P-piste):
P et P1 étant deux pistes dont les traces P et P1 sont à priori indépendantes.
- Une P-piste associée à P1, notée P.P1, est l'ensemble formé des candidats de P, ceux de P1 et des candidats $A_i \in G$ que l’on placerait avec les TR si les candidats de P et de P1 étaient placés.
- Si P1 est issue d’un candidat A_k , on écrira P.P(A_k).
- Si P1 est issue d’un ensemble de candidats E, on écrira P.P(E).
Lorsque P1 est une antipiste, alors, P.P1 est une P-antipiste, et,
- Si P1 est issue d’un candidat A_k, on écrira P.P'(A_k).
- Si P1 est issue d’un ensemble de candidats E, écrira P.P'(E).
La trace d’une P-piste P.P1 est la P-piste P.P1 obtenue en utilisant seulement les TB ou les TE.

A noter, que si P1 est incluse dans P, la P-piste est identique à P. Pour engendrer une P-piste qui n'est pas réduite à P, ce qui serait sans intérêt puisque P est à ce stade bloquée dans sa construction par les TB, P1 ne doit pas à priori être incluse dans P.
Ainsi par exemple, A_k ne doit pas être choisi parmi les candidats connus de P pour générer une P-piste P.P(A_k). Cela n'exclu pas que, ultérieurement et en fonction des avancées obtenues dans la résolution, A_k puisse devenir un candidat de P une fois la piste P totalement construite.

Pour illustrer cette définition voici, sur la figure suivante, un exemple de construction partielle d’une P-piste.

La piste P marquée en bleu issue du 2L1C5 est construite avec les TB seulement, sa trace P ne compte que deux candidats.
La piste P1 issue du 1L5C1 est construite avec les TB seulement, sa trace P1 ne compte que 4 candidats 1L5C1, 2L4C1, 1L1C2 et 7L4C2.
La P-piste P.P(1L5C1) (du moins sa trace) est composée de tous les candidats marqués en jaune et bleu.

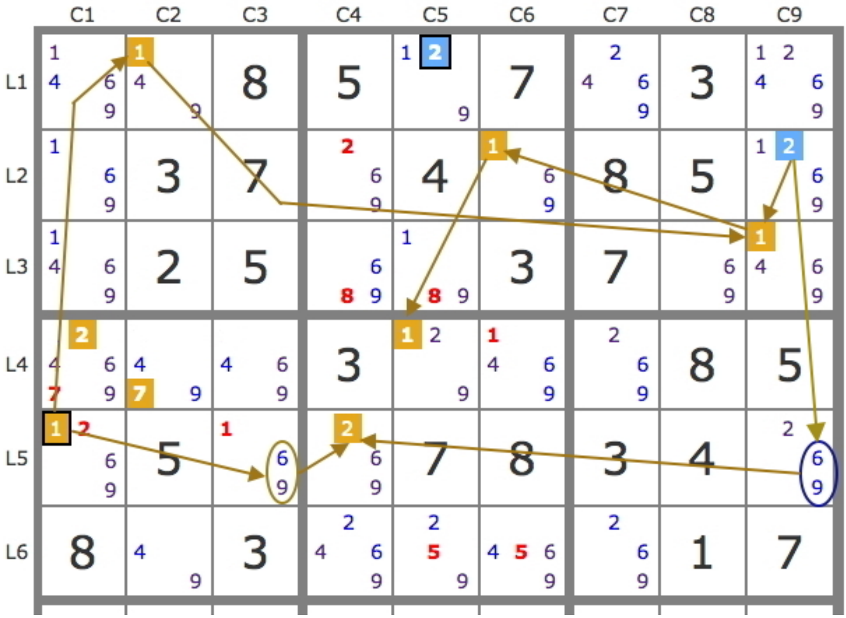

Fig 4 : P(2L1C5) et P-piste P.P(1L5C1)

On peut énoncer pour les P-pistes des propriétés et définitions similaires à celles des chapitres précédents, notamment les suivantes.

Propriété 8-1 :
a) Si P1 ⊆ P2, alors P.P1 ⊆ P.P2.
b) Si P1 ⊆ P.P2, alors P.P1 ⊆ P.P2.

En effet :
a) si un candidat A_i est placé par les TR en supposant que les candidats de P et P1 sont placés, à fortiori ce candidat est placé par les TR en supposant que les candidats de P et P2 sont placés puisque les candidats de P1 sont placés avec ceux de P2. Donc tout candidat de P.P1 étant un candidat de P.P2, on a P.P1 ⊆ P.P2.
b) si un candidat A_i est placé par les TR en supposant que les candidats de P et P1 sont placés, à fortiori ce candidat est placé par les TR en supposant que les candidats de P et P.P2 sont placés puisque les candidats de P1 sont placés avec ceux de P.P2. Donc tout candidat de P.P1 étant un candidat de P.(P.P2), on a P.P1 ⊆ P.(P.P2), soit P.P1 ⊆ P.P2 puisque (voir annexe) P.(P.P2)= P.P2.

Propriété 8-2 :
Si E1 $\subseteq$ E2, alors P.P(E2) $\subseteq$ P.P(E1).

En effet, d'après la propriété 2-2 on a P(E2) $\subseteq$ P(E1), donc d'après la propriété 8-1 on a P.P(E2) $\subseteq$ P.P(E1).

Propriété 8-3 :
Pour toute P-piste P.P1,
$\forall A_i \in$ P.P1 on a : $P(A_i) \subseteq$ P.P1 et $P.P(A_i) \subseteq$ P.P1.

En effet, si $A_i \in$ P.P1 c'est que A_i a été placé par les TR en supposant que les candidats de P et de P1 sont placés. En plaçant un candidat A_k par les TR comme un candidat de $P(A_i)$ on ne fait que poursuivre, par définition, la construction de P.P1. Donc $A_k \in$ P.P1 aussi. Tout candidat de $P(A_i)$ est donc un candidat de P.P1, soit $P(A_i) \subseteq$ P.P1.
Dès lors $P.P(A_i) \subseteq$ P.P1.

Définitions 8-2 (validité, invalidité) :
1) Une P-piste P.P1 est valide, si et seulement si, tous les candidats de P et de P1 sont des candidats solutions de la même solution.
Une P-piste P.P1 est S_n-valide, si et seulement si, P.P1 $\subseteq S_n$.
2) Une P-piste P.P1 est invalide lorsque le placement des candidats qui la composent conduit à une incompatibilité avec R.
Une P-piste P.P1 est S_n-invalide, si et seulement si, P.P1 $\not\subseteq S_n$.

Propriété 8-4 :
Pour toute P-piste P.P1 on peut énoncer que :
a) Une P-piste P.P1 est valide, si et seulement si, $\exists S_n$ telle que P.P1 $\subseteq S_n$.
b) P.P1 est invalide, si et seulement si, $\forall S_n$ P.P1 est S_n-invalide.

En effet,
a) si P.P1 est valide, P et P1 ne sont composées que de candidats solutions d'une même solution S_n , donc, par définition de P.P1, P.P1 n'est composée que de candidats solution $\in S_n$ aussi, c'est-à-dire que P.P1$\subseteq S_n$ et P.P1 est S_n-valide.
Inversement, si $\exists S_n$ telle que P.P1 est S_n-valide alors P $\subseteq$ P.P1 $\subseteq S_n$ et P1 $\subseteq$ P.P1 $\subseteq S_n$ donc P et P1 ne sont composées que de candidats solutions de la même solution, P.P1 est donc valide par définition.

b) si P.P1 est invalide et si pour une S_n on avait P.P1 $\subseteq S_n$, tous les candidats de P.P1 seraient des candidats solutions $\in S_n$ ce qui est en contradiction avec le fait que leur placement est incompatible avec R.

Donc, si une piste est invalide on a $P \not\subset S_n$ $\forall S_n$, c'est-à-dire que P est S_n-invalide $\forall S_n$.
Inversement, si $\forall S_n$ P.P1 est S_n-invalide, alors $P.P1 \not\subset S_n$ $\forall S_n$ et P.P1 ne peut pas être valide car, ce serait une contradiction, P.P1 valide entraînerait qu'il existe au moins une S_n telle que $P.P1 \subseteq S_n$. Donc P.P1 est invalide.

Propriété 8-5 :
Soient deux P-pistes P.P1 et P.P2 telles que $P.P1 \subseteq P.P2$.
a) Si P.P2 est S_n-valide, alors P.P1 est S_n-valide.
b) Si P.P1 est S_n-invalide, alors P.P2 est S_n-invalide.
c) Si P.P1 est invalide, alors P.P2 est invalide.

En effet,
a) Si P.P2 est S_n-valide on a $P.P1 \subseteq P.P2 \subseteq S_n$ donc P.P1 est S_n-valide.
b) Si P.P1 est S_n-invalide on a $P.P1 \not\subset S_n$ donc $P.P2 \not\subset S_n$ et P.P2 est S_n-invalide.
c) Si P.P1 est invalide, alors P.P1 est S_n-invalide $\forall S_n$ donc P.P2 est S_n-invalide $\forall S_n$, c'est-à-dire que P.P2 est invalide.

Ces définitions et propriétés étant acquises, nous pouvons définir ce qu'est une extension d'une piste P, par la définition suivante :

Définition 8-3 (extension d'une piste) :
La P-piste P(E)=P.P(E) est appellée une extension de P issue de E lorsque la P-piste P.P'(E) est invalide .*
Les P-pistes P.P(E_i), où $E = \cup_i E_i$, sont les branches de P.*
On dit aussi que les branches de P forment une bifurcation de P.*

A noter, évidemment, que la validité et l'invalidité de P* sont définies par celles de P.P(E) selon la définition 8-2 et la propriété 8-4.
Etablissons alors les propriétés suivantes.

Propriété 8-6 :
Pour toute extension P de P, on peut énoncer que:*
a) Si P est S_n-valide, alors P est S_n-valide et une branche de P* est Sn-valide.*
b) Si toutes les branches de P sont S_n-invalides, alors P est S_n-invalide.*
c) Si P est S_n-invalide, alors P est S_n-invalide.*

En effet :
a) Si P est S_n-valide, tous les candidats de P sont des candidats $\in S_n$.

Dès lors, P.P'(E) étant invalide, si aucun $A_k \in E$ n'était un candidat solution $\in S_n$, alors par définition de P.P'(E) tous les candidats de P.P'(E) seraient des candidats solutions $\in S_n$ ce qui est en contradiction avec le fait que leur placement est incompatible avec R. Donc un $A_k \in E$ au moins est un candidat solution $\in S_n$ et par définition de P.P(A_k) tous les candidats de P.P(A_k) sont des candidats solutions $\in S_n$, c'est-à-dire que P.P(A_k) $\subseteq S_n$ et une branche (cette branche) de P* est S_n-valide. Enfin P.P(E) est aussi S_n-valide puisque P.P(E) $\subseteq$ P.P(A_k).
b) Cette assertion découle de la précédente a), puisque dans le cas contraire, si P était S_n-valide alors P* serait S_n-valide et P* $\subseteq S_n$ ce qui est en contradiction du fait que, P* étant S_n-invalide $\forall S_n$ P* $\not\subset S_n$.
c) Cette assertion découle de la précédente b) d'après la propriété 8-5 car P.P(E) $\subseteq$ P.P(E_k) $\forall$k.

Le théorème suivant trouve une application immédiate dans la résolution des grilles en facilitant le prolongement d'une piste arrêtée à sa trace.

Théorème 8-1 :
E1 et E2 formant une paire d'ensembles cachée de la trace $\underline{P}$ de P,
Si P.P(E2) est invalide, alors P.P(E1) est une extension de P.

En effet : P.P'(E1) est construite en supposant que les candidats de E1 sont éliminés, donc que seuls les candidats de E2 subsistent comme candidats pouvant être des candidats de P.P'(E1) puisque E1 et E2 forment une paire d'ensembles caché (voire annexe) de $\underline{P}$. Comme un des candidats $A_k \in$ E2 doit être placé pour appliquer R on peut dire que un des $A_k \in$ E2 est un candidat de P.P'(E1), donc que pour ce A_k on a (d'après la propriété 8-3) P.P(A_k) $\subseteq$ P.P'(E1), donc que P.P(E2) $\subseteq$ P.P(A_k) $\subseteq$ P.P'(E1) puisque P(E2) $\subseteq$ P(A_k) entraîne P.P(E2) $\subseteq$ P.P(A_k).
Si P.P(E2) est invalide, alors d'après la propriété 8-5 P.P'(E1) est invalide donc P.P(E1) est une extension de P par définition.

En particulier, A1 et A2 formant une paire cachée de candidats, si P.P(A2) est invalide, P.P(A1) est une extension de $\underline{P}$.

Etablissons enfin les propriétés suivantes utiles à la résolution des grilles.

Propriété 8-7 (dualité) :
P1 et P2 étant des pistes conjuguées, si P1 est invalide, tous les candidats de P2* sont des candidats solutions communs à toutes les solutions de la grille. .*

En effet, si P1* est invalide alors P1* est S_n-invalide $\forall S_n$, donc P1 est S_n-invalide $\forall S_n$. Si P2 est une piste conjuguée de P1 alors P2 est S_n-valide $\forall S_n$, donc P2* est S_n-valide $\forall S_n$ aussi d'après la propriété 8-6, c'est-à-dire que ses candidats sont des candidats solutions communs à toutes les solutions.

Pour établir la deuxième propriété, donnons la définition suivante.

Définitions 8-4:
On dit qu'un candidat voit une extension P lorsqu'il voit un candidat faisant partie de l'extension P*.*

Propriété 8-8 (élimination) :
P1 et P2 étant des pistes conjuguées, un candidat qui voit P1 et P2* n'est pas un candidat solution de S_n , $\forall S_n$, et peut être éliminé de la grille.*

Soit A_n un candidat qui voit un candidat A_i de P1*=P.P(E1) et un candidat A_r de P2*=P.P(E2).
Si A_n était un candidat solution $\in S_n$, A_i et A_r ne seraient pas des candidats $\in S_n$ en vertu de R, donc P(A_i) et P(A_r) seraient toutes deux S_n-invalides, soit P($A_\mathbf{i}$) $\not\subset S_n$ et P($A_\mathbf{i}$) $\not\subset S_n$.
Comme d'après la propriété 8-3 P(A_i) $\subseteq$ P.P(E1) et P(A_r) $\subseteq$ P.P(E2), on aurait P1* $\not\subset S_n$ et P2* $\not\subset S_n$, et donc P1* et P2* serait S_n-invalides ainsi que P1 et P2 d'après la propriété 8-6, ce qui est impossible pour des pistes conjuguées.
Donc A_n n'est pas un candidat solution de S_n et ce $\forall S_n$, il peut donc être éliminé de la grille.

Propriété 8-9 (validation) :
P1 et P2 étant des pistes conjuguées, un candidat commun à P1 et P2* est un candidat solution commun à toutes les solutions de la grille.*

En effet, cette propriété est une conséquence de la propriété 8-8, puisque tous les autres candidats de la case contenant un candidat commun à P1* et P2* voient P1* et P2*, et en conséqence peuvent être éliminés $\forall S_n$. Ne subsiste donc plus dans cette case comme candidat pouvant être solution que le candidat commun à P1* et P2*.

Propriété 8-10:
Si E est une entité cachée de P, P.P(E) est une extension de P.

En effet, si E est une entité cachée de P (voir annexe), P.P'(E) est invalide puisque tous les candidats de E étant éliminés, R ne peut plus être respectée dans l'entité contenant E.

9. Application

Voici deux exemples d'utilisation d'extensions d'une piste permettant de débloquer une grille.
Reprenons la grille de la figure 3 pour la traiter autrement.

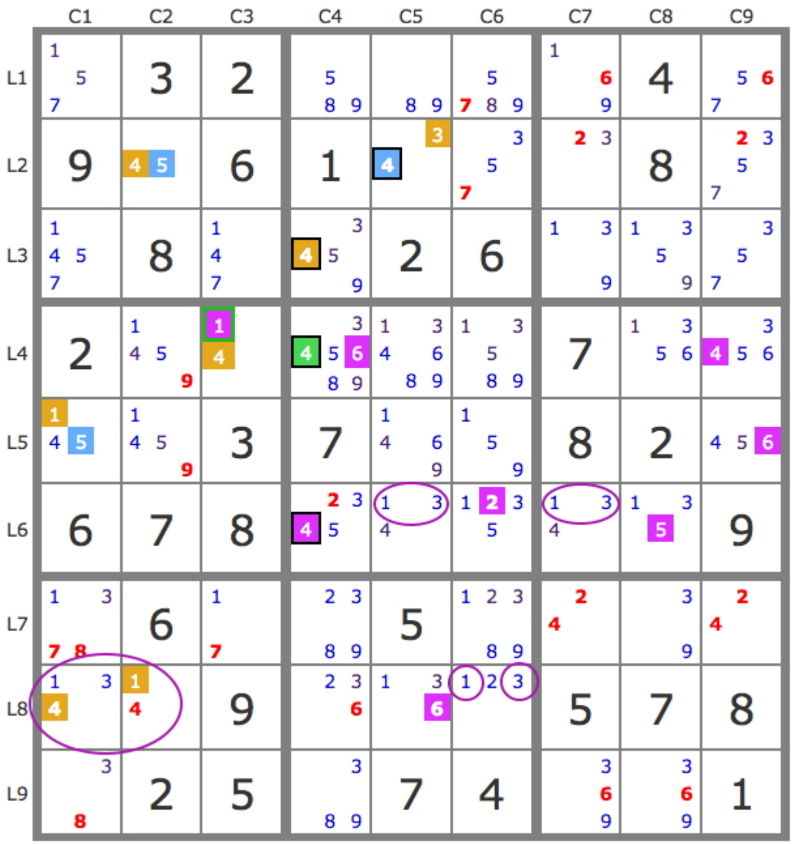

Fig 5 : Extension d'une piste

La trace de la piste bleue P ne comptant que 3 candidats, on utilise une extension P*=P.P(4L46C4) dont les deux composantes sont les P-pistes verte issue du 4L4C4 et violette issue du 4L6C4 (figure 5). P*=P.P(4L46C4) est bien une extension de P, puisque la P-piste 4L46C4 est une entité cachée de P (propriété 8-10).
Les deux branches de P* ayant en commun le 1L4C3 celui-ci est donc un candidat de P*, laquelle se développe alors très significativement par les TB au point de couvrir la grille pour fournir une solution. On vérifiera aussi que la piste jaune est invalide.

Une autre façon de traiter cette grille est d'utiliser le jeu de pistes conjuguées issues de la paire 14L4C3 (figure 6).

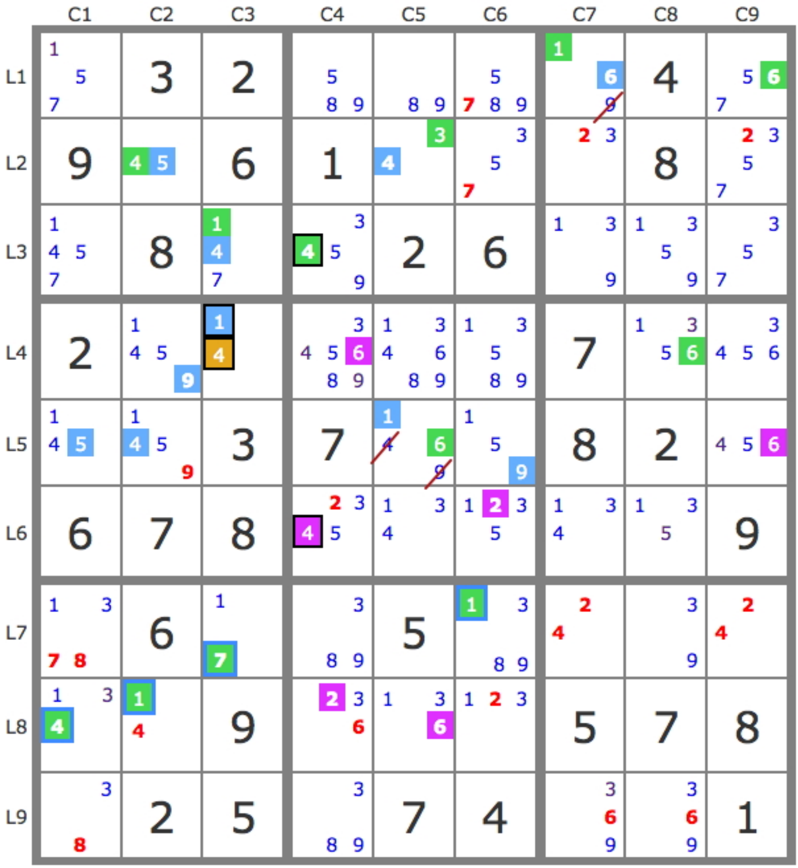

Fig 6 : résolution d'une grille

La piste bleue P1 issue du 1L4C3 est développée partiellement. La piste jaune P2 est réduite à sa trace qui compte un seul candidat. Comme la P-piste violette P2.P(4L6C4) est invalide car elle ne peut pas avoir de candidat 4 dans B6, la P-piste verte P2.P(4L3C4) est une extension P2* de la piste jaune qui a en commun avec la piste bleue P1 (donc P1*) les candidats 4L8C1, 1L8C2 et 7L7C3, lesquels sont donc des candidats solutions de la grille (propriété 8-9). Certains candidats (barrés en rouge) peuvent être éliminés puisqu'ils voient P1 (donc P1*) et P2* (propriété 8-8).
En regardant de plus près, on constate aussi que P2* est invalide (pas de candidat vert dans L8C5), donc que tous les candidats de P1 sont des candidats solutions (propriété 8-7).

10. Résolution d'une grille.

La résolution d'une grille revient à exploiter les propriétés d'un jeu de pistes conjuguées pour déteminer, au moins partiellement, des candidats solutions, puis si nécessaire d'utiliser d'autres jeux de pistes jusqu'à résoudre complètement la grille. C'est ce qu'on appelle une résolution par jeux de pistes successifs.

Pour les grilles très difficiles ce mode de résolution ne suffit pas en raison de la difficulté que l'on peut avoir à construire les pistes et leurs extensions. Pour certaines pistes une extension ne suffit pas, aussi est-on amené à utiliser des extensions de branches d'extensions, c'est à dire des P-pistes de P-pistes.

Définitions 10-1 (Cascade de P-pistes) :
Une cascade de P-pistes est une suite P_0, P_1, P_2,… P_i …P_k où P_0 =P et P_i (i>0) est la P-piste P_{i-1}.P(E_i) (ou la P-antipiste P_{i-1}.P'(E_i)) définie par la définition 8-1 dans laquelle on remplaçe P par P_{i-1}.

La figure 7 est une représentation d'une cascade de P-pistes.

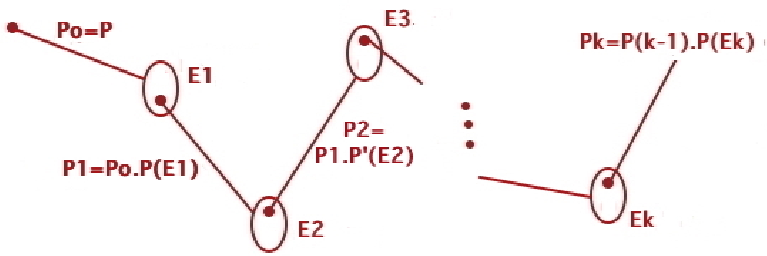

Fig 7 : cascade de P-pistes

Définitions 10-2 (validité) :
Une cascade de P-piste P_0, P_1,… P_i …P_k est S_n-valide, si et seulement si, $P_i \subseteq S_n$ $\forall$ i.

Définitions 10-3 (Arbre d'extensions) :
- Un arbre d'extensions d'une piste P est un ensemble P^*_1, P^*_2,… P^*_i …P^*_m où P^*_1 est une extension de P et ou chaque P^*_i (i>1) est l'extension d'une branche B_{kr} d'une autre extension P^*_k avec P^*_i=B_{kr}.P(E_i) et B_{kr}.P'(E_i) invalide (voir définition 8-3).
- L'extension d'une branche est dite placée à l'extrémité de la branche et prolonge la branche.
- Les branches qui ne sont pas prolongées par une extension, sont appelées les branches extrêmes (ou extrémités) de l'arbre.
- Une cascade de branches est une cascade de P-pistes où les P-pistes sont des branches.

La figure 8, où chaque rond rouge représente une extension et chaque segment rouge représente une branche d'extension, est une représentation d'un arbre d'extensions.
La chaîne de segments verts représente une cascade de branches.

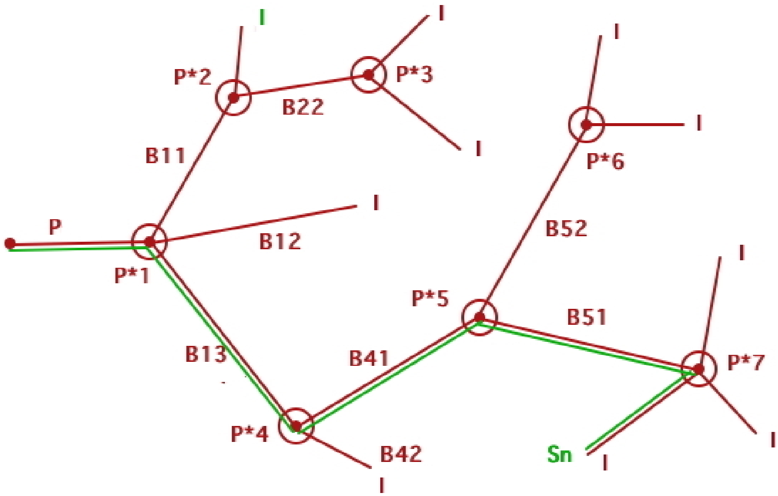

Fig 8 : arbre d'extensions d'une piste

Propriété 10-1 :
Soit P^*_1, P^*_2,… P^*_m un arbre d'extensions de P.
a) Si P est S_n-valide, alors une cascade de branches est S_n-valide, c'est-à-dire que chaque branche de la cascade est S_n-valide .
b) Si toutes les branches extrêmes de l'arbre sont invalides, alors P est invalide.

En effet,
a) pour k=1, P^*_1 est l'extension de P et si P est S_n-valide une branche de P^*_1 est S_n-valide aussi d'après la propriété 8-6.
L'assertion étant vraie à l'ordre k=1, supposons la vraie à l'ordre k (k<m), c'est-à-dire vraie pour une cascade de k branches qui sont donc S_n-valides si P est S_n-valide. Désignons par B_k la dernière branche de la cascade et P^*_{k+1} l'extension située à son extrémité. Comme $B_k.P'(E_{k+1})$ est invalide par définition de P^*_{k+1} , on peut dire que si aucun $A_r \in E_{k+1}$ n'est un candidat solution $\in S_n$, alors par définition de $B_k.P'(E_{k+1})$ tous les candidats de $B_k.P'(E_{k+1})$ sont des candidats solutions $\in S_n$ ce qui est en contradiction avec le fait que leur placement est incompatible avec R.
Donc un $A_r \in E_{k+1}=\cup_i E_{(k+1)i}$ au moins est un candidat solution $\in S_n$, d'où $A_r \in E_{(k+1)i}$ pour un $E_{(k+1)i}$ au moins et pour ce $E_{(k+1)i}$, par définition de $B_k.P(E_{(k+1)i})$ tous les candidats de $B_k.P(E_{(k+1)i})$ sont des candidats solutions $\in S_n$, c'est-à-dire que $B_k.P(E_{(k+1)i}) \subseteq S_n$ et donc qu'une branche de P^*_k est S_n-valide. Ce qui démontre l'assertion pour tout k.

b) Toutes les branches extrêmes étant invalides, si P était valide, il existerait S_n telle que P est S_n-valide et il existerait une cascade de branches S_n-valides dont tous les candidats de la dernière branche B_r, qui est une branche extrême, sont des candidats solutions $\in S_n$ ce qui est en contradiction avec le fait que leur placement est incompatible avec R. Donc P est invalide, ce qui démontre l'assertion.

Terminons ce chapitre par la définition suivante :

Définitions 10-4 (Arbre de résolution) :
Un arbre de résolution est l'arbre formé des arbres d'extensions respectifs de deux pistes conjuguées.

Il découle de cette définition que si un des deux arbres d'extensions à toutes ses extrémités invalides, l'autre contient une cascade de branches valides dont tous les candidats sont solutions de la grille.

11. Unicité et niveau de difficulté TDP

La technique des pistes permet d'évaluer le niveau de difficulté d'une grille à solution unique différemment de celui, dit conventionnel, obtenu par les techniques évoluées (X-wing, Swordfich, ALS, …), cela en apportant la preuve que la grille, dont à priori rien n'indique qu'elle n'a qu'une seule solution, est bien à solution unique.

Propriété 11-1 :
Si *pour chaque jeu de pistes utilisé lors de la résolution de la grille par la technique des pistes, le placement des candidats solutions est obtenu par application des propriétés d'interactions 4-2,3,4, 6-1-1, 8-7,8,9 et 10-1 alors la grille a une solution unique.*

En effet, selon ces propriétés, tout candidat placé est commun à toutes les solutions S_n. Ainsi, si on aboutit par ce procédé à construire une solution, toutes les solutions ont les mêmes candidats, c'est à dire qu'il n'y a qu'une solution.
Par exemple, la solution de la grille de la figure 5 est unique puisque pour sa construction on a utilisé l'interaction des deux branches de l'extension de la piste bleue et que la piste jaune conjuguée de la piste bleue est invalide.

Définition 11-1 :
La taille d'une résolution d'une grille à solution unique est le nombre d'invalidités, autres que celles liées aux entités cachées, nécessaires pour établir l'unicité de la solution lors de la construction de cette solution par la technique des pistes utilisant les seules TB.

Par exemple, si la solution d'une grille à solution unique est obtenue par l'utilisation de deux jeux des pistes conjuguées, le premier jeu en constatant l'invalidité d'une des deux pistes, le second jeu en constatant l'invalidité des deux branches d'une extension d'une des deux pistes, la taille de la résolution est de 3 : +1 pour l'invalidité du premier jeu de pistes et +2 pour l'invalidité des deux branches de l'extension utilisée dans le second jeu de pistes.

Cette définition de la taille a l'inconvénient de poser problème dans le cas où un jeu de pistes conjuguées est formé de deux pistes valides (ce qui est rare mais possible) puisqu'alors il n'y a pas d'invalidité associée à ce jeu de pistes. C'est aussi le cas pour deux pistes ou deux extensions dont seul le croisement (propriétés 4-4 et 8-9) est utilisé.
C'est pourquoi il convient de complèter cette définition de la taille en précisant qu'il faut compter les invalidités sous-jacentes pour chaque

pistes ou extension dont l'invalidité d'une des deux pistes (ou branches) n'a pas été utilisée ou établie.

Par exemple, une résolution réalisée avec un seul jeu de pistes conjuguées par croisement (propriété 4-4) ne fait apparaître aucune invalidité. Sa taille n'est pas zéro mais 1.

La taille d'une résolution peut être calculée autrement par la formule suivante qui ne fait intervenir que le nombre de jeux de pistes et des branches des arbres de résolution utilisées pour résoudre une grille et montrer l'unicité de sa solution.

Théorème 11-1 (calcul de la taille)(*)
J_2 étant le nombre de jeux de pistes conjuguées et d'extensions à deux branches, et J_k pour $k>2$ étant le nombre d'extensions à k branches, la taille d'une résolution est égale à :
$T= J_2 + 2J_3 + \ldots + (k-1) J_k + \ldots + (n-1) J_n$.

En effet, la résolution d'une grille est décrite comme sur la figure ci-dessous par des arbres de résolution associés aux jeux de pistes conjuguées JP.

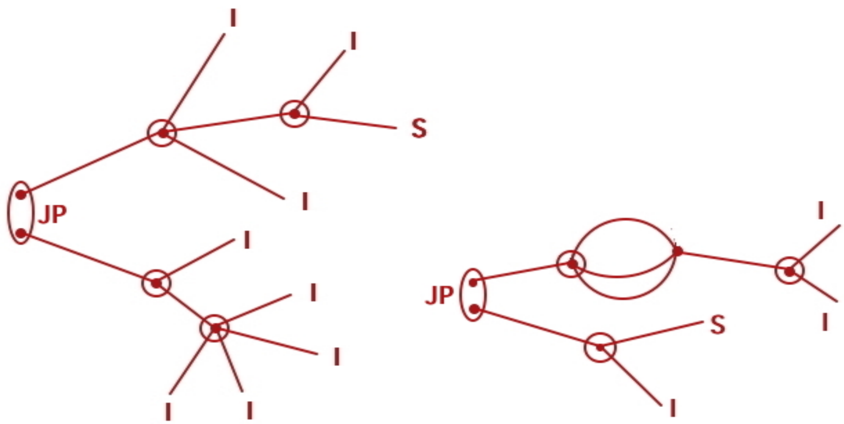

Fig 9 : arbres de résolution

Traitons d'abord le cas d'un seul jeu de pistes (figure de gauche). La taille T de la résolution est déterminée par le nombre d'extrémités invalides. Ainsi, si N est le nombre d'extrémités de l'arbre, on a T=N-1 puisqu'il faut soustraire l'extrémité associée à la solution (S).

La formule donnant T est vérifiée pour une résolution réalisée avec le seul jeu de pistes JP, puisque la taille de cette résolution est T=1 et que J_2 =1 et J_k= 0 ∀k>2.
Supposons alors que la formule est vérifiée pour une résolution réalisée par un jeu de pistes JP et par $(J_2-1)+ J_3 \ldots + J_k \ldots + J_n$ extensions totalisant $2(J_2-1)+ 3J_3 \ldots+ 4J_k \ldots + nJ_n$ branches à N extrémités.
Si une résolution est réalisée avec un jeu de pistes nécessitant une extension supplémentaire de m branches prolongeant une extrémité, le nombre d'extrémités de l'arbre de résolution devient N'= N + m-1 et la taille de cette résolution devient T' = N'-1=N+m-2= T+m-1, soit :
- si m<n+1, $T' = J_2 + 2J_3 +\ldots+ (m-1)J'_m+\ldots + (n-1) J_n$, où $J'_{m=} J_m+1$.
- si m >n, $T' = J_2 + 2J_3 + \ldots + (n-1) J_n + (m-1) J_m$, où J_m=1.
La formule est encore vérifiée à l'ordre suivant, elle est donc vérifiée à tous ordres.

Pour tenir compte de la situation particulière où les branches d'une extension convergent vers un candidat ou ensemble de candidats permettant de développer l'extension sans mettre en évidence les invalidités de ses branches (figure de droite de la figure 9), il faut dans cette formule compter cette extension pour 1 dans J_p si elle a p branches.

Enfin, si la résolution d'une grille demande plusieurs jeux de pistes JP_1, JP_2, … JP_r assortis de leurs bifurcations donnant chacun un calcul partiel T1, T2, etc… Tr, alors T=T1+T2+… +Tr satisfait la formule puisque :
$T = J_2 + 2J_3 + \ldots (k-1) J_k\ldots+ (q-1) J_q$ où $J_{k=} J_{k1}+ J_{k2}+\ldots+ J_{kr}$ ∀k,
q étant le plus grand nombre de branches d'une même extension.

A titre d'exemple d'utilisation de cette formule, pour la résolution représentée sur la figure de gauche de la figure 9 on a T=8, tandis que pour celle représentée sur la figure de droite on a T= 5, et pour les deux réunies on a T=13.
Dans les résolutions présentées au chapitre 9, T=2 dans les deux cas.

Définition 11-2 :
Le niveau TDP d'une grille est la taille minimale de toutes les résolutions possibles de la grille par la technique des pistes.
On convient dans cette définition que le niveau TDP = 0 est celui d'une grille que l'on peut résoudre par les seules TB.

Cette définition n'est pas facile à mettre en œuvre car il n'est pas concevable de trouver toutes les résolutions possibles par la technique des pistes. Elle a toutefois l'avantage de fixer une borne supérieure qui est la plus petite taille obtenue parmi les résolutions réalisées et ainsi donner une assez bonne estimation de la difficulté d'une grille.

Voici une comparaison empirique sur la base de nombreuses grilles traitées, entre les niveaux conventionnels et les niveaux TDP. Cela donne une bonne idée des niveaux de difficultés TDP des grilles usuelles et une bonne indication du nombre de jeux de pistes et/ou extensions à utiliser pour les résoudre.

Niveau conventionnel	Niveau TDP
1 à 6	0
7 à 9	1
10 à 12	2
13 à 16	3
17 à 20	4

12. Conclusion

Ce qui caractérise la Technique des Pistes, c'est que seules les TB sont nécessaires pour résoudre une grille puisque, pour la trace d'une piste comme pour l'extension ou un arbre de résolution d'une piste, les TB suffisent pour construire une solution.

Evidemment rien n'interdit, d'utiliser des techniques élémentaires (TB, X-wing, etc…) dans le cadre de la technique des pistes au lieu des seules TB.

D'un point de vue pratique, la technique des pistes est présentée et illustrée par de nombreux exemples dans mon livre "Technique des pistes en sudoku" et sur le site internet qui l'accompagne www.assistant-sudoku.com .

Les grilles proposées au public sont généralement à solution unique. Pour s'exercer à la technique des pistes sur des grilles à solutions multiples, on pourra utiliser les Grilles N°359 et N°425 du site internet www.assistant-sudoku.com , et plus généralement sur une grille à solution unique de laquelle on retire un candidat dévoilé.

14. Annexe

On présente dans cette annexe, quelques résultats et définitions nécessaires aux démonstrations du chapitre 8 et 9.

- Comme l'ordre importe peu pour placer les candidats avec TR lorsqu'on place les candidats de P puis ceux de P1 , ou les candidats de P1 puis ceux de P, on a **P.P1=P1.P**

- **(P.P1).P =P.P1= P.(P.P1).**
En effet, les candidats de P.P1 étant placés, de placer P à nouveau n'ajoute aucun candidat supplémentaire à P.P1.

- Une entité cachée de P (trace de P) est une entité de laquelle on retire les candidats qui voient P.

- Deux ensembles disjoints E1 et E2 forment une paire d'ensembles cachée de P lorsque leur réunion est une entité cachée de P.

(*) Remerciements

Ce document consacré à une présentation qui se veut rigoureuse de la technique des pistes n'aurait pas été aussi abouti sans les remarques judicieuses de certains utilisateurs de la technique que je remercie ici, et parmi lesquels M. François Cordoliani qui a fait une relecture critique de toutes les démonstrations.
On doit à Monsieur François Cordoliani l'énoncé et la démonstration du thèorème 11-1.

Toulouse octobre 2015, mise à jour novembre 2016, janvier 2017, décembre 2017.