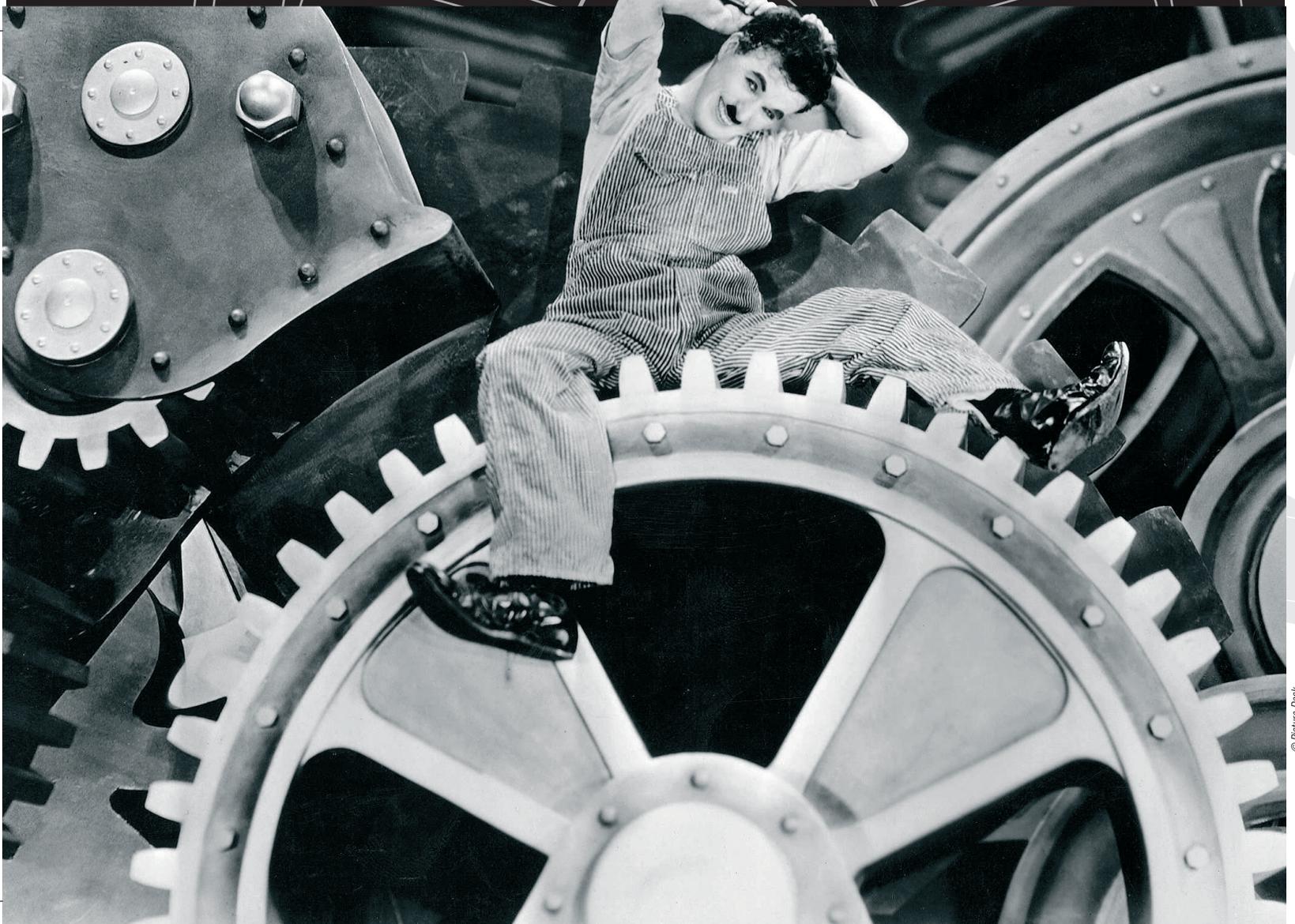


# L'algèbre des sandwiches

*Pour modéliser et analyser des systèmes tels que chaînes de production ou réseaux de transport, les mathématiciens remplacent l'addition et la multiplication par des opérations plus exotiques.*

Guy Cohen • Stéphane Gaubert • Jean-Pierre Quadrat



**N**ous avons tous appris à l'école que l'on n'additionne pas des oranges et des bananes. Ne pourrions-nous ajouter que des quantités de même nature, exprimées dans une même unité? L'expression « un mètre plus deux kilogrammes » serait-elle vide de sens? Non. Il existe des situations de la vie courante où nous transgressons ces règles sans y penser et où nous utilisons une autre façon de compter. Un fabricant d'automobiles n'a aucun mal à « additionner » un moteur, une carrosserie et quatre roues : cela donne une voiture. Pour l'employé d'une boulangerie qui fabrique des sandwiches jambon-fromage, un morceau de baguette « plus » une tranche de jambon « plus » une lamelle de fromage font un sandwich. Dans cette situation, l'addition de trois quantités de natures différentes donne un résultat de nature distincte. Selon cette opération, «  $1 + 1 + 1 = 1$  ». De surcroît, si l'employé garnit le sandwich de deux tranches de jambon au lieu d'une, le résultat reste un sandwich unique. Autrement dit, «  $2 + 1 + 1 = 1$  ». L'opération revient à prendre le minimum des quantités qu'il faut associer : le nombre de sandwiches préparés est égal au minimum des quantités de pain, de jambon et de fromage disponibles.

Cette étrange façon de compter correspond à une structure mathématique nommée algèbre *min-plus*. Ce n'est pas une curiosité mathématique : elle permet, avec l'algèbre associée *max-plus*, de décrire la plupart des systèmes qui nécessitent une synchronisation de ressources, tels les ateliers de production industriels, les correspondances d'un réseau ferroviaire ou la gestion des feux de circulation. Après avoir présenté ces systèmes dits à événements discrets et leur représentation graphique, nous détaillerons les propriétés des algèbres *min-plus* et *max-plus*, puis nous analyserons ces systèmes à l'aide de ces outils mathématiques.

Dans l'étude des systèmes dynamiques classiques, on examine l'évolution d'un état du système en fonction du temps, qui peut s'écouler de façon continue ou progresser par à-coups, mais de façon autonome. Or certains systèmes dynamiques fonctionnent différemment. Imaginons un voyageur qui désire aller en train dans une ville voisine. Qu'il se rende à la gare une demi-heure ou cinq minutes en avance par rapport à l'horaire de départ ne changera pas son heure d'arrivée : il doit patienter jusqu'à l'entrée en gare du train. Durant le trajet, rien ne se passe jusqu'à ce qu'il descende et prenne une correspondance. Les seuls événements significatifs de cette situation sont les rencontres du voyageur avec le train, puis avec la correspondance. L'état de ce type de système n'évolue que lors de ces instants ponctuels définis par

leur dynamique même, indépendamment de l'écoulement du temps. On les nomme systèmes à événements discrets.

Dans ces systèmes, les actions ne peuvent être accomplies que lorsqu'un certain nombre de conditions (autorisations ou ressources matérielles) sont réunies. L'enchaînement des processus dans nombre de systèmes réels – réseaux de transport, chaînes de production, automates – repose sur cette synchronisation : un voyage nécessite la conjonction d'un voyageur sur le quai et d'un train arrivant en gare ; la préparation d'un sandwich réclame simultanément du pain, du jambon et du fromage ; l'assemblage d'une automobile n'est possible que si une machine et des pièces sont disponibles en même temps, etc.

Alors que l'évolution des systèmes dynamiques naturels obéissant aux lois de la physique est décrite par des équations différentielles classiques, les systèmes à événements discrets ne se laissent pas modéliser facilement par les outils de l'analyse mathématique. L'emploi des algèbres *max-plus* et *min-plus* permet de mettre ces phénomènes de synchronisation en équations. Pour cela, nous partirons d'une représentation graphique fondée sur les *réseaux de Petri*.

## Places, jetons et transitions

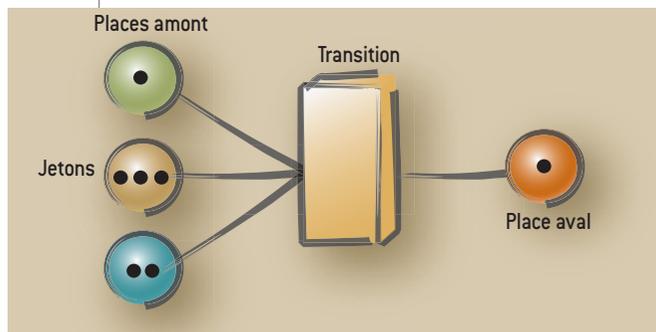
Les réseaux de Petri ont été introduits en 1962 par le mathématicien Carl Petri, alors à l'Université de Bonn, dans le cadre de la théorie des automates. Ce sont des graphes – des figures formées de lignes, les arêtes, et de points de jonction de ces lignes, les nœuds – qui représentent les systèmes à événements discrets.

Les réseaux de Petri comportent deux sortes de nœuds : les transitions et les places. Chaque arête relie une transition à une place. Le sens de parcours des arêtes, imposé, définit les places amont et aval d'une transition. Des jetons éventuellement présents dans les places constituent un marquage du réseau qui caractérise l'état du système représenté ; ce marquage évolue dans le temps. Lorsque chacune des places en amont d'une transition contient au moins un jeton, celle-ci est activée. Cette activation entraîne le retrait d'un jeton dans les places amont et l'ajout d'un jeton dans l'ensemble des places aval. L'ensemble des transitions du réseau de Petri réunissant les conditions nécessaires est simultanément activé.

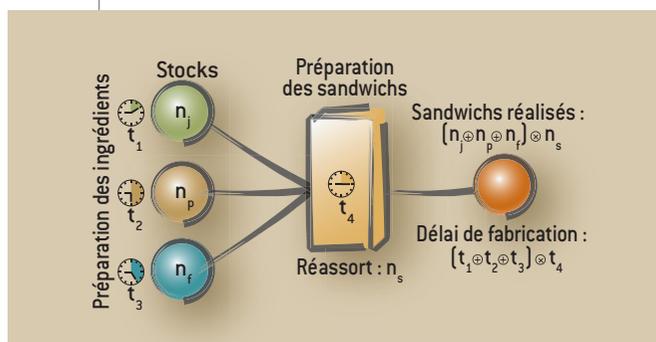
Un tel fonctionnement reproduit celui des systèmes régis par la disponibilité synchronisée des ressources. Les transitions symbolisent les actions, et les jetons matérialisent à la fois les ressources nécessaires à la réalisation des actions et le produit de ces actions. Les places représentent les stocks de ressources ou de produits. La réalisation de sandwiches précédemment évoquée est ainsi modélisée par une transition précédée de trois places et suivie d'une autre place (voir l'encadré 2). On réalise des modèles graphiques

**1. Charlie Chaplin** dans *Les temps modernes* (1936) caricature la volonté « moderne » de gagner du temps et l'organisation qui en résulte.

## 2. Le graphe du sandwich



Un réseau de Petri modélise le fonctionnement des systèmes synchronisés, ici la réalisation de sandwiches. La transition (*rectangle*) représente l'action d'assemblage des ingrédients. Les places amont (*ronds à gauche*) symbolisent les ressources nécessaires (pain, jambon et fromage). Le nombre de jetons présents dans ces places correspond aux quantités d'ingrédients disponibles. Le marquage de la place aval (*à droite*) indique le nombre de sandwiches confectionnés et mis à disposition des clients. Lorsque chaque ingrédient est disponible (il y a au moins un jeton dans chaque place amont), l'employé assemble un sandwich (la transition est activée), puis le met en vente (ajout d'un jeton dans la place aval), tandis que les stocks de chaque ingrédient diminuent (disparition d'un jeton dans chacune des places amont).



Le nombre de sandwiches réalisables dépend des stocks. Une tranche de jambon, trois de fromages, et deux morceaux de pain limitent au minimum, ici un, le nombre de sandwiches réalisables. Si l'employé réapprovisionne les stocks avec deux parts de chaque ingrédient, le nombre de sandwiches réalisables passe à trois.  $n_j$ ,  $n_p$  et  $n_f$  désignant les stocks initiaux de jambon, de pain et de fromage et  $n_s$  le réapprovisionnement, le nombre de sandwiches réalisables s'exprime dans le formalisme de l'algèbre *min-plus* par l'équation  $n = (n_j \oplus n_p \oplus n_f) \otimes n_s$ . L'équation permet de calculer les stocks initiaux minimaux nécessaires à la conception de  $n$  sandwiches.

À l'inverse, estimer le délai de fabrication du sandwich en fonction des dates de disponibilité des ingrédients  $t_1$ ,  $t_2$  et  $t_3$  et de la durée de préparation  $t_4$  met en jeu l'algèbre *max-plus*. Le délai de fabrication  $d$  est égal au maximum des temps  $t_1$ ,  $t_2$  et  $t_3$  auquel s'ajoute  $t_4$ , c'est-à-dire, dans l'algèbre *max-plus*,  $d = (t_1 \oplus t_2 \oplus t_3) \otimes t_4$ . L'équation permet de déterminer les dates les plus tardives de livraison des ingrédients de telle sorte que les sandwiches soient mis en vente à une date donnée.

du fonctionnement de systèmes complexes en assemblant plusieurs sous-ensembles constitués d'une transition précédée et suivie d'une ou plusieurs places. Les biens produits en aval d'une transition jouent alors le rôle de ressources pour les transitions suivantes. Dans le modèle de la boulangerie, la place aval peut approvisionner des transitions représentant la vente d'un sandwich à un client.

Les réseaux de Petri ont d'abord été utilisés dans un contexte où l'écoulement du temps n'intervient pas. Les logiciens et les informaticiens les employaient pour vérifier qu'aucune situation de blocage ne pouvait apparaître dans des chaînes d'implications logiques représentant le déroulement d'une séquence d'instructions. Dans le domaine de la productique, nous cherchons en revanche à déterminer l'évolution des systèmes physiques, notamment la quantité de biens produite par unité de temps. Nous décrivons cette activité par un modèle en *compteurs*, en établissant des équations où chaque variable exprime le nombre d'activations cumulées d'une transition du réseau en fonction du temps.

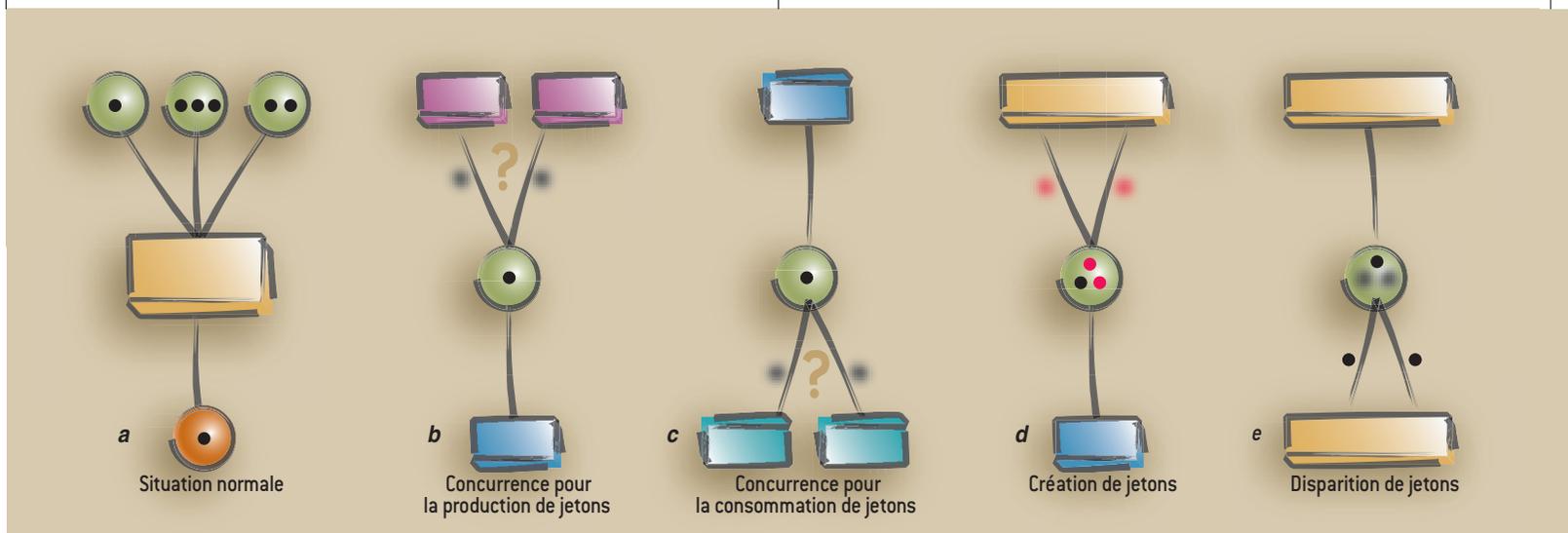
Symétriquement, dans un modèle en *dateurs*, les variables des équations expriment le temps écoulé en fonction du nombre d'activations des transitions. Ce modèle permet de calculer les temps de production et de traiter des problèmes comme la gestion des stocks à « flux tendus », gestion qui consiste à acheter les matières premières le plus tard possible et à limiter le stockage tout en respectant les dates de livraison des produits.

## Technique japonaise

La modélisation de la gestion à flux tendus par des réseaux de Petri est analogue à une méthode de production japonaise mise en œuvre dans les années 1960, le *Kanban* (étiquette). À l'entrée d'un atelier, des étiquettes disponibles sur un tableau sont collées sur les pièces. Elles sont récupérées lors de la sortie de l'atelier et remises sur le tableau d'entrée. De nouvelles pièces ne sont insérées dans la chaîne de fabrication que si des étiquettes sont disponibles, c'est-à-dire si l'état du stock de sortie le nécessite.

L'étude de la dynamique des systèmes synchronisés nécessite ainsi l'introduction du temps dans les réseaux de Petri dits alors temporisés. Dans les systèmes réels, l'immobilisation d'une pièce dans un magasin ou le trajet d'un train entre deux correspondances nécessitent une durée minimale. Le séjour d'un jeton dans une place du réseau est ainsi affecté d'une durée. De même, comme la réalisation d'une opération prend un certain temps, une durée est attribuée à l'activation des transitions. Il est cependant équivalent et plus commode de modéliser les transitions temporisées par deux transitions instantanées de début et de fin encadrant une place portant la durée de l'action. La transition « fabriquer un sandwich » se décompose par exemple en une transition « commencer la réalisation », suivie d'une place représentant les 30 secondes nécessaires à la préparation, puis d'une transition « fin de la préparation ». Dans les réseaux de Petri temporisés, les durées sont ainsi portées par les places et correspondent à un temps de séjour minimal des jetons.

La circulation des jetons de place en place dans un réseau de Petri est parfois problématique. Lorsque deux transitions



**3. Un réseau de Petri** est formé de places (les cercles) et de transitions (les rectangles) reliées par des arêtes. Des jetons disposés dans les places représentent des ressources ou des autorisations disponibles. Une transition est activée lorsque toutes les places amont contiennent au moins un jeton. Cela consomme un jeton dans chaque place amont, et en produit un dans chaque place aval. De nouveaux biens ou autorisations sont ainsi disponibles en aval du graphe. Les réseaux dont les places possèdent une seule arête entrante et une sortante, et où un seul jeton rentre ou sort à chaque activation, sont nommés graphes d'événements (a). Ils sont

déterministes, viables et modélisables par des équations linéaires dans des algèbres appropriées. Des problèmes apparaissent lorsque cette configuration n'est pas respectée. Si un même bien est créé par deux actions différentes, la linéarité est compromise (b). Lorsque la même ressource est partagée par deux actions distinctes (c), l'indétermination nécessite un choix extérieur. La présence d'une transition amont délivrant plusieurs jetons simultanément conduit à l'augmentation exponentielle du nombre de jetons (c), tandis qu'une transition aval consommant plusieurs jetons mène à la mort du réseau par la disparition des ressources (d).

se partageant la même place amont, par exemple, laquelle est activée lorsqu'un jeton est présent dans cette place? Cette configuration entraîne un conflit entre deux actions pour l'attribution d'une ressource unique. Le système n'est pas correctement défini. Il existe de nombreux cas menant à des problèmes de nature diverse. Les réseaux de Petri qui en sont exempts sont nommés graphes d'événements (voir la figure 3). Chaque place n'y possède qu'une seule arête incidente et une seule arête sortante, et la transition amont ne produit qu'un jeton tandis que la transition aval n'en consomme qu'un à chaque activation. Cette configuration a pour conséquence la conservation du nombre de jetons dans chaque circuit du graphe au cours du temps, et donc la viabilité du réseau dès lors que chaque circuit comporte initialement au moins un jeton. La dynamique d'un tel système est alors entièrement définie par le marquage initial du réseau et par une seule règle: activer les transitions dès que les ressources sont disponibles.

Le principal attrait des graphes d'événements est l'existence de techniques algébriques permettant de décrire leur dynamique. Au prix d'un changement de structure algébrique, les équations aux dateurs ou aux compteurs s'expriment sous la forme d'équations récurrentes (l'état du système dépend des états précédents) et surtout linéaires (qui ne comportent que des sommes de variables pondérées par des coefficients multiplicatifs connus), dont la résolution est bien connue.

Pour ce faire, il suffit de compter différemment: les graphes d'événements et les systèmes à événements discrets sont intimement liés à la façon inhabituelle de calculer évoquée précédemment. Dans ces situations, les opérations consistant à prendre le maximum ou le minimum de plusieurs nombres jouent un rôle clé. L'activation d'une transition dans un graphe d'événements requiert et consomme un jeton dans chacune des places amont. Elle peut être répétée jusqu'à ce qu'une de ces places soit vide, soit un nombre de fois égal au minimum des quantités de jetons contenus dans les places amont. Inversement, la date d'activation d'une transition est dictée par la date de disponibilité la plus tardive, donc maximale, des jetons dans les places amont. De même, dans un atelier, la production de biens est limitée par la ressource en plus

faible quantité, et la date de réalisation d'un objet ne peut être inférieure au plus long des temps de mise à disposition de ses constituants. En d'autres termes, compter des quantités produites revient à additionner les ressources en prenant le minimum ( $1 + 2 + 3 = 1$ ), tandis que calculer des temps de production revient à additionner les délais de fabrication des composantes ou d'acheminement des ressources en prenant le maximum ( $1 + 2 + 3 = 3$ ).

L'addition habituelle est également déterminante dans ces situations. Dans un modèle en dateurs, la date de disponibilité d'un jeton est la somme de sa date d'arrivée dans la place et de la durée d'attente associée, tandis qu'en compteurs, la quantité de jetons disponibles dans une place résulte de l'ajout de la quantité initiale et de celle produite par les activations successives de la transition.

## Des algèbres exotiques

Les algèbres *max-plus* et *min-plus*, dont la structure est adaptée aux opérations propres à ces situations – maximum ou minimum et addition –, sont le cadre naturel pour décrire les systèmes à événements discrets. Détaillons leur construction. L'addition et la multiplication ordinaires vérifient de nombreuses propriétés. En particulier, pour chacune de ces opérations, il existe un nombre qui n'a pas d'effet, appelé élément neutre: 0 pour l'addition et 1 pour la multiplication. Les deux opérations sont associatives: on peut regrouper les calculs de plusieurs façons sans modifier le résultat. De surcroît, l'addition est commutative – l'ordre de calcul n'importe pas – tandis que la multiplication est distributive par rapport à l'addition:  $(a + b) \times c = a \times c + b \times c$ . Ces propriétés font de l'ensemble des nombres réels muni de ces deux opérations une structure algébrique nommée semi-anneau. L'addition et la multiplication sont les opérations les plus naturelles, mais tout couple d'opérations qui vérifie les mêmes propriétés forme un semi-anneau.

Remplaçons l'addition par l'opération consistant à prendre le maximum de plusieurs nombres, opération notée *max*, et la multiplication par l'addition habituelle. Ces nouvelles opérations vérifient les propriétés précédentes. Elles

## 4. Les chemins optimaux et l'algèbre des matrices

Les matrices carrées sont des tableaux de nombres comportant autant de lignes que de colonnes. Dans le cadre des algèbres *max-plus* ou *min-plus*, nous verrons comment déterminer avec ces matrices des chemins optimaux dans des graphes.

La position d'un élément dans une matrice est indiquée par deux indices,  $i$  pour la ligne et  $j$  pour la colonne. Dans la matrice  $A$ , l'élément en position  $(1,1)$ , noté  $a_{11}$ , vaut 2.

$$A = \begin{pmatrix} 2 & 5 \\ 3 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$$

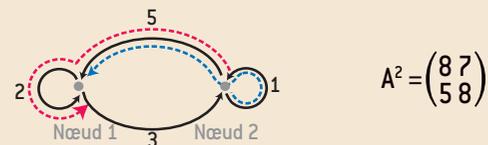
Tout comme les nombres, on peut additionner et multiplier les matrices. Chaque élément de la matrice «  $A + B$  » est la somme des éléments correspondants des matrices  $A$  et  $B$ . Un élément de position  $(i, j)$  du produit  $AB$  est la somme des produits terme à terme des éléments de la ligne  $i$  de la matrice  $A$ , et de la colonne  $j$  de la matrice  $B$ .

$$A + B = \begin{pmatrix} 3 & 7 \\ 7 & 4 \end{pmatrix} \quad A \times B = \begin{pmatrix} 2 & 5 \\ 3 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix} = \begin{pmatrix} 22 & 19 \\ 7 & 9 \end{pmatrix}$$

Dans l'algèbre *max-plus*, le résultat de ces opérations change. L'élément de position  $(1, 2)$  de la matrice  $A \oplus B$  est ainsi le maximum des éléments  $a_{12}$  et  $b_{12}$  dans  $A$  et  $B$  :  $(A \oplus B)_{12} = 5 \oplus 2 = 5$ . De même, l'élément de position  $(1, 2)$  du produit  $A \otimes B$  est le maximum des nombres obtenus en sommant terme à terme la ligne et la colonne encadrées :  $(A \otimes B)_{12} = (2 \otimes 2) \oplus (5 \otimes 3) = \max(2 + 2, 5 + 3) = 8$ .

$$A \otimes B = \begin{pmatrix} 2 & 5 \\ 3 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 5 \\ 3 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix} = \begin{pmatrix} 8 & 8 \\ 5 & 5 \end{pmatrix}$$

À chaque matrice carrée correspond un graphe. Il comporte autant de nœuds que la matrice compte de colonnes (et de lignes), et une arête joint les nœuds  $j$  et  $i$  si l'élément  $a_{ij}$  de position  $(i, j)$  de la matrice  $A$  n'est pas nul. Ce nombre définit le « poids » de cette arête, par exemple, dans un cadre ferroviaire, la durée d'un parcours ou la longueur d'un trajet. L'élément  $a_{11}$  de la matrice  $A$  précédente étant égal à 2, par exemple, une boucle de poids 2 relie le nœud 1 à lui-même. Le graphe associé à la matrice  $A$  est ainsi :



Le produit de matrices dans l'algèbre *max-plus* s'interprète en termes de chemins optimaux dans un tel graphe. Tandis que les éléments de la matrice  $A$  sont les poids des chemins de longueur un, ceux de la matrice  $A^2$  sont les poids maximaux des chemins à deux arêtes joignant deux nœuds. L'élément de position  $(i, j)$  de la matrice  $A^2 = A \otimes A$  est en effet la somme (dans *max-plus*) des produits  $a_{ik} \otimes a_{kj}$  pour toutes les valeurs de  $k$ . Le terme  $a_{ik} \otimes a_{kj} = a_{ik} + a_{kj}$  donne le poids du

sont évidemment associatives. La première est commutative, car le maximum de deux nombres ne dépend pas de leur ordre, et la deuxième est distributive par rapport à la première :  $(\max(a, b)) + c = \max(a + c, b + c)$ . Adjoindre l'infini négatif,  $-\infty$ , à l'ensemble des réels permet de doter l'opération *max* d'un élément neutre. En effet,  $\max(-\infty, a) = a$  pour tout nombre  $a$ . L'élément neutre de la deuxième opération est bien sûr celui de l'addition ordinaire, le zéro. Le maximum et l'addition forment ainsi une structure de semi-anneau sur l'ensemble des réels. On la nomme algèbre *max-plus*. Par analogie, on note la première loi  $\oplus$  et la deuxième loi  $\otimes$ .

L'algèbre *max-plus* a des propriétés inhabituelles. L'addition d'un nombre avec lui-même redonne le même nombre puisque l'opération revient à prendre le maximum de deux nombres égaux. L'addition est dite idempotente et elle n'est pas simplifiable, c'est-à-dire que lorsque  $a \oplus b = a \oplus c$ , on ne peut pas conclure que  $b = c$ . En outre, les nombres « négatifs » par rapport au « zéro » de l'algèbre *max-plus* n'existent pas : un nombre  $n$  a pas d'opposé par rapport à  $-\infty$  (pour tout nombre  $a$ , il n'existe pas de nombre  $b$  tel que  $\max(a, b) = -\infty$ ). En d'autres termes, il n'y a pas de soustraction dans l'algèbre *max-plus*.

Cette absence semble présager de difficultés pour résoudre des équations dans l'algèbre *max-plus*. Une propriété pallie cependant en partie l'absence de signe moins. L'égalité «  $a \oplus b = a$  », c'est-à-dire  $\max(a, b) = a$ , implique que  $b$  soit inférieur à  $a$ . La réciproque est vraie. Cette relation d'ordre découlant de l'idempotence de l'addition permet de classer les nombres réels selon l'ordre naturel et de résoudre, au moins de manière approchée, des équations mettant en jeu des fonctions croissantes.

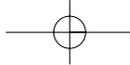
De façon similaire, on construit l'algèbre *min-plus*. Le minimum – noté également  $\oplus$  – remplace dans ce cas

l'addition usuelle, laquelle fait office de multiplication. L'algèbre *min-plus* jouit des mêmes propriétés que sa jumelle *max-plus*, à ceci près que l'ordre découlant de l'idempotence est l'inverse de l'ordre naturel («  $a \oplus b = a$  », c'est-à-dire  $\min(a, b) = a$ , implique dans cette algèbre que  $a$  est supérieur à  $b$ , ce qui est contraire à l'ordre usuel). Les algèbres *max-plus* et *min-plus* sont étroitement liées. En particulier, l'application qui transforme un nombre en son opposé (au sens usuel) associée à un élément de l'algèbre *max-plus* un unique nombre dans l'algèbre *min-plus* (et réciproquement) tout en respectant les lois de ces algèbres : on parle alors d'isomorphisme. Nous verrons que cet isomorphisme, associé à l'échange des coefficients multiplicatifs avec les termes de « décalage », permet de passer des équations du modèle en dateurs dans *max-plus* aux équations du modèle en compteurs dans *min-plus*.

### L'équation des sandwiches

Les opérations *max* et *min* sont *a priori* peu sympathiques pour les mathématiciens : elles ne sont ni simplifiables ni dérivables. Pour autant, elles offrent une description mathématique des systèmes à événements discrets en termes d'équations linéaires. Celles-ci se prêtent, nous allons le voir, à l'emploi de techniques algébriques tel le calcul matriciel adapté au cadre de ces algèbres.

La production d'un poste de travail et les délais à respecter, par exemple ceux d'un boulanger concoquant des sandwiches (représenté par une transition et des places amont et aval) s'expriment par des équations linéaires dans les algèbres *max-plus* et *min-plus* (voir l'encadré 2). La description de l'évolution temporelle d'un système à événements discrets mène ainsi à un ensemble d'équations linéaires récurrentes



chemin formé des arêtes successives de  $j$  à  $k$  et de  $k$  à  $i$ . L'élément de position  $(i, j)$  de  $A^2$  est donc le poids du chemin de poids maximal parmi tous ceux joignant les nœuds  $j$  et  $i$  en deux étapes. Dans le graphe précédent, il y a, par exemple, deux chemins menant du nœud 2 au nœud 1. Le premier,  $2 \rightarrow 1 \rightarrow 1$ , a pour poids  $5 + 2 = 7$  (*chemin rouge*), et le deuxième,  $2 \rightarrow 2 \rightarrow 1$ , vaut  $1 + 5 = 6$  (*chemin bleu*). L'élément de position  $(1, 2)$  de  $A^2$ , maximum de 6 et 7, vaut donc 7.

De manière générale, l'élément de position  $(i, j)$  de la matrice  $A^n$  fournit le poids maximal parmi les chemins de  $n$  étapes joignant les nœuds  $j$  et  $i$ . Cette propriété a des applications en théorie des jeux. Imaginons un joueur dont les décisions successives amènent des gains divers. La situation est modélisée par des déplacements dans un graphe : le joueur reçoit la somme  $a_{ij}$  lorsqu'il passe de  $j$  à  $i$ . Son revenu optimal en  $n$  déplacements est donné par la puissance  $n$ -ième dans *max-plus* de la matrice associée au graphe. L'algèbre *max-plus* est ainsi « l'algèbre des chemins de gain maximal » dans les graphes. Dans l'algèbre *min-plus*, on obtient le décompte des chemins les moins coûteux.

L'algèbre *max-min*, avec le maximum comme addition, et le minimum comme multiplication, est aussi utile dans les problèmes de cheminement. L'élément  $a_{ij}$  symbolise le débit d'un tuyau reliant  $j$  à  $i$ . Le débit d'une canalisation étant limité par le minimum des débits des tuyaux qui la composent, calculer  $A^n$  dans l'algèbre *max-min* revient à trouver les canalisations de débit maximal parmi celles de longueur  $n$ . On qualifie parfois l'algèbre *max-min* d'algèbre « goulot ».

couplées. Examinons le graphe d'événements de la sandwicherie qui comporte des transitions symbolisant la fabrication, la vente, le réapprovisionnement ou l'arrivée des clients. Chaque place porte une durée : délai d'approvisionnement, temps de réalisation d'un sandwich, repos de l'employé, etc. Le taux de production de ce système, qui décrit le nombre de sandwiches fabriqués en fonction du temps, mais aussi, par exemple, le taux de renouvellement des stocks, est donné par le nombre d'activations de chaque transition par unité de temps. Cela s'exprime par une équation dans l'algèbre *min-plus*, car  $c$  est la disponibilité de la ressource la plus rare qui déclenche les activations.

Le nombre d'activations à l'instant  $t$  de la transition  $x_i$ , noté  $x_i(t)$ , est égal au minimum des nombres d'activations des transitions amont  $x_j(t - t_{ij})$ ,  $x_k(t - t_{ik})$ , ..., où  $t_{ij}$ , par exemple, désigne la durée de stationnement minimale dans la place située entre les transitions  $x_j$  et  $x_i$ , auxquels s'ajoutent les stocks initiaux  $s_{ij}$ ,  $s_{ik}$ , ..., de jetons dans ces places. Pour qu'une transition soit en effet activée à un moment donné, il faut que les transitions amont aient été activées un moment auparavant, ce qui assure l'approvisionnement des places intermédiaires. Le décalage dépend de la durée d'attente dans ces places et intervient comme un facteur de retard dans l'équation. Les stocks de jetons initiaux dans ces mêmes places augmentent potentiellement le nombre minimal d'activations de la transition et interviennent comme des termes additifs (au sens habituel). Le nombre d'activations de la transition est égal au minimum des ressources disponibles dans les places amont  $x_j$ ,  $x_k$ , ..., c'est-à-dire  $x_i(t) = \min(x_j(t - t_{ij}) + s_{ij}, x_k(t - t_{ik}) + s_{ik}, \dots)$ . Dans l'algèbre *min-plus*, le minimum se note  $\oplus$ , tandis que les termes additifs deviennent des coefficients multiplicatifs ; l'équation s'écrit donc :  $x_i(t) = s_{ij} \otimes x_j(t - t_{ij}) \oplus s_{ik} \otimes x_k(t - t_{ik}) + \dots$  Elle est linéaire.

L'ensemble des équations pour chaque transition dans le modèle en compteurs décrit l'évolution du système à événements discrets. On résume ces informations dans un tableau de nombres – une matrice  $A$  – dont les éléments (appartenant à l'algèbre *min-plus*) représentent les coefficients des variables dans le système d'équations. Cette matrice ne dépend que du marquage initial du graphe d'événements. Le système s'écrit après quelques manipulations sous forme d'une équation matricielle récurrente qui exprime l'état  $X(t)$  du système – un vecteur dont les coordonnées représentent le nombre d'activations de l'ensemble des transitions au temps  $t$  – en fonction de son état à l'instant précédent  $X(t - 1)$ . Pour passer d'un état au suivant, il suffit de lui appliquer la matrice du système :  $X(t) = A X(t - 1)$ .

La détermination des équations du modèle en dateurs du même système, qui décrivent les dates d'activation de chaque transition et expriment donc les délais de production, conduit de façon semblable à un système d'équations linéaires dans l'algèbre *max-plus*. Les modèles faisant appel aux deux algèbres sont duaux. La temporisation des places, qui donne des termes de retard dans les équations aux compteurs dans *min-plus*, engendre des coefficients multiplicatifs dans les équations aux dateurs dans *max-plus*, tandis que les stocks donnant des coefficients multiplicatifs deviennent des termes qui décalent les dates d'activations. Les équations  $x(k) = 2 \otimes x(k - 1)$  dans l'algèbre *max-plus* et  $x(t) = 1 \otimes x(t - 2)$  dans l'algèbre *min-plus*, par exemple, décrivent le même système physique. Les ressources initiales et les temporisations des places sont des quantités duales, ce qui résulte de l'isomorphisme entre les deux algèbres et des fonctions de dateurs ou de compteurs qui sont inverses l'une de l'autre.

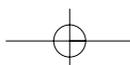
## Une approche analytique

À l'équation matricielle de base se greffent souvent des termes de « contrôle », qui modélisent les données externes régulant le fonctionnement du système à événements discrets. En l'absence de clients, par exemple, les sandwiches ne s'écoulent pas, tandis qu'avec une file d'attente infinie, les sandwiches trouvent toujours acquéreur. Une transition correspondant à l'arrivée des clients représente ainsi une donnée externe. C'est par le biais de telles entrées que les concepteurs régulent des systèmes synchronisés, en délivrant de façon maîtrisée des ressources dans des places en amont des transitions à contrôler.

Sous la forme matricielle, la dynamique des systèmes à événements discrets se prête à une approche analytique. Ce point de vue a permis de comprendre des comportements qualitatifs comme l'existence de régimes périodiques et de goulots d'étranglement dans les réseaux, inexplicables par les seules simulations informatiques.

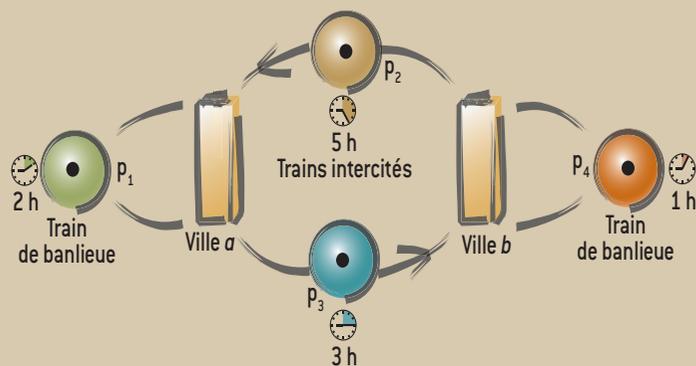
En appliquant de façon répétée la matrice décrivant un système à un vecteur représentant l'état de ce système, on calcule les états successifs à partir d'une configuration initiale donnée. On est donc amené à calculer les puissances de la matrice  $A$ . Cette opération  $a$ , dans les algèbres *min-plus* ou *max-plus*, des propriétés surprenantes qui s'interprètent en termes de chemins dans le graphe d'événements.

Si l'on définit le poids d'un chemin comme le nombre total de jetons présents dans les places traversées, le produit



## 5. Correspondances périodiques

Illustrons les techniques d'analyse qu'autorisent les algèbres *min-plus* et *max-plus* par un modèle de réseau de transport. Deux villes *a* et *b* sont reliées par des trains. Le trajet de *a* vers *b* dure 3 heures, alors que le retour sur une autre voie, en travaux, dure 5 heures. Chacune des deux villes est desservie par un train de banlieue. Le premier décrit une boucle qui le ramène à la gare de la ville *a* toutes les 2 heures, tandis que le deuxième passe toutes les heures en gare dans la ville *b*.



Graphe d'événements d'un réseau ferroviaire

Ce réseau de transport est un système à événements discrets : les correspondances dans chaque ville ne sont possibles que lorsque le train de banlieue local et le train en provenance de l'autre ville sont tous deux arrivés. Étudions la dynamique de ce réseau, c'est-à-dire le nombre de correspondances réalisées en fonction du temps et du nombre de trains en circulation, et estimons son efficacité.

Une brève observation du système montre que, s'il y a un train sur chacune des voies, c'est le circuit aller-retour reliant les deux villes qui limite les performances du réseau : l'aller-retour prenant 8 heures et ce circuit comportant deux arrêts, son rendement est de quatre heures par arrêt. Dans le cas d'un réseau plus complexe, il peut cependant être très difficile, voire impossible, de déterminer empiriquement le circuit limitant et le régime périodique. Le recours au formalisme de l'algèbre *max-plus* est alors indispensable. Détaillons-en la méthode.

Nous modélisons le réseau de transport par un graphe d'événements comportant deux transitions *a* et *b* correspondant aux deux villes, deux arêtes orientées entre ces transitions portant des places  $p_2$  et  $p_3$  affectées de temporisations (les pondérations du cas général) de 5 et 3 unités respectivement, plus des boucles à chaque transition munies de places  $p_1$  et  $p_4$  portant des temporisations de 2 et 1 respectivement.

Le nombre de correspondances en fonction du temps est décrit par des équations aux dateurs ; nous nous plaçons donc dans le cadre de l'algèbre *max-plus*. Pour qu'une correspondance se produise dans la ville *a*, il faut qu'un train parte de la ville *b* 5 heures plus tôt, et donc que la correspondance précédente y ait eu lieu 5 heures auparavant (la récurrence dépend du nombre de trains sur la voie. S'il y en avait deux, la *n*-ième correspondance dans la ville *a* dépendrait de la (*n*-2)-ième correspondance dans la ville *b*). De même, il faut qu'il y ait eu une correspondance du train de banlieue dans la ville *a* 2 heures auparavant. La date à laquelle la *n*-ième correspondance dans la ville *a*, notée  $a(n)$ , peut se réaliser est donc celle du plus tardif de ces événements, c'est-à-dire le maximum des dates  $b(n-1) + 5$  et  $a(n-1) + 2$ .

Dans le langage des graphes d'événements, la transition *a* est activée quand les places  $p_1$  et  $p_2$  contiennent des jetons qui y ont séjourné respectivement 2 et 5 heures. Ces places reçoivent elles-mêmes un jeton après l'activation des transitions en amont. La *n*-ième activation de *a* prend place 2 heures après la précédente dans la même gare, ou 5 heures après la (*n*-1)-ième activation de *b*. La date de cet événement est donc égale au maximum de ces deux dates. La temporisation des places intervient comme un terme additif (dans *max-plus*) dans la mise en équations, tandis que les stocks de ressources initiaux agissent sur le degré de récurrence de l'équation.

Dans l'algèbre *max-plus*, le maximum et l'addition remplacent la somme et le produit, de sorte que l'équation aux dateurs de *a* s'écrit :

$$a(n) = 5 \otimes b(n-1) \oplus 2 \otimes a(n-1).$$

De même, l'équation pour *b* est :

$$b(n) = 3 \otimes a(n-1) \oplus 1 \otimes b(n-1).$$

Ces équations sont linéaires, de sorte que l'on peut écrire ce système de deux équations sous forme matricielle.

$$\begin{pmatrix} a \\ b \end{pmatrix} (n) = \begin{pmatrix} 2 & 5 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} (n-1)$$

La matrice *A* caractérise le système :

$$A = \begin{pmatrix} 2 & 5 \\ 3 & 1 \end{pmatrix}$$

En appliquant successivement cette matrice au vecteur d'état initial du système  $X(0) = (a(0), b(0))$ , on calcule l'évolution des dates de correspondances entre les trains, et on détermine le régime périodique du système.

$$\text{Ainsi, } X(n) = A.X(n-1) = A^2.X(n-2) = \dots = A^n.X(0)$$

On calcule les puissances successives de *A* en multipliant lignes par colonnes les termes de *A*, avec les règles de calcul de l'algèbre *max-plus*. On obtient :

$$A^2 = \begin{pmatrix} 8 & 7 \\ 5 & 8 \end{pmatrix}, A^3 = \begin{pmatrix} 10 & 13 \\ 11 & 10 \end{pmatrix}, A^4 = \begin{pmatrix} 16 & 15 \\ 13 & 16 \end{pmatrix} = \begin{pmatrix} 8+8, & 7+8 \\ 5+8, & 8+8 \end{pmatrix} = 8 \otimes A^2$$

En poursuivant, on voit que  $A^5 = 8 \otimes A^3$ , etc. Après quatre itérations, le système se cale sur un régime à deux périodes alternées de six et deux heures. Les horaires des correspondances dans les deux villes progressent ainsi alternativement de deux et six heures et toutes les huit heures passent exactement deux trains.

Cette valeur qui caractérise la matrice est aussi liée au cycle limitant les performances du système. La temporisation totale du cycle reliant les deux villes est en effet de huit heures, pour seulement deux jetons ; le rapport vaut quatre. Cela signifie que deux trains se partageant huit heures de trajet effectuent une correspondance toutes les quatre heures en moyenne. En comparaison, les circuits de banlieue sont plus efficaces : le ratio entre les temporisations et les ressources y vaut deux et un respectivement. Pour améliorer le rendement du réseau, il faut accélérer le circuit inter-cités, en y rajoutant deux trains afin que le ratio descende à deux.

La modélisation dans l'algèbre *min-plus* permet de compter le nombre de correspondances en fonction du temps (équations aux compteurs). Pour la ville *a*, par exemple, on obtient l'équation :

$$a(t) = 1 \otimes b(t-3) \oplus 1 \otimes a(t-2).$$

Les temporisations deviennent des retards et les ressources des termes multiplicatifs : les deux modélisations sont duales.

matriciel dans l'algèbre *min-plus* (modélisation en compteurs) revient à sélectionner les chemins les plus « légers » dans le graphe d'événements. Dans l'algèbre *max-plus* (modèle en dateurs), le poids est donné de façon similaire par la durée affectée à chaque place et le produit matriciel correspond à la sélection des chemins les plus « lourds » (voir l'encadré 5). La matrice  $A$  correspond ainsi aux chemins de longueur un,  $A^2$  désigne les chemins de longueur deux, etc. Lorsque l'on élève la matrice  $A$  à une puissance excédant sa taille, la longueur des chemins parcourus dépasse le nombre de transitions du graphe et les chemins repassent par des transitions déjà visitées : ils incluent des cycles. Le poids des chemins suffisamment longs est minimisé (dans *min-plus*) lorsque ceux-ci passent de manière répétée par le cycle de poids minimum. À partir d'un rang très élevé, les éléments des puissances successives de la matrice croissent donc de façon périodique selon un taux égal au poids moyen de ce cycle minimal (son poids divisé par sa longueur).

## Cycles et valeurs propres

Dans les années 1980, lors de simulations de fonctionnement d'ateliers de production, nous avons découvert un phénomène analogue. Après un laps de temps variable, les transitions d'un graphe d'événements s'activent à intervalles réguliers de façon synchronisée. Le système passe d'un mode transitoire à un régime périodique. Ce comportement, entravé par les boucles de contrôle ajoutées par les concepteurs, n'apparaissait pas clairement et nous ignorions s'il était intrinsèque aux systèmes à événements discrets. Leur modélisation algébrique a révélé que cette périodicité était liée à des quantités caractéristiques du système, les valeurs propres.

La notion de valeur propre est essentielle à l'analyse des systèmes dynamiques. La note émise par une corde vibrante, le balancement périodique d'un pendule ou le mode de vibration d'une cloche sont déterminés par les valeurs propres de la matrice décrivant le système. Une valeur propre est un nombre associé à un vecteur dit propre sur lequel l'action de la matrice équivaut à une multiplication par la valeur propre. Dans l'algèbre *min-plus*, où la multiplication est l'addition usuelle, l'action de la matrice sur le vecteur propre revient à additionner la valeur propre à toutes les composantes du vecteur propre. En d'autres termes, quand le réseau est dans un état correspondant à un vecteur propre de la matrice, le passage à l'état suivant, sous l'action de la matrice, consiste à activer toutes les transitions un nombre de fois égal à la valeur propre associée dans le modèle en compteurs, ou à décaler toutes les dates de la même quantité dans le modèle en dateurs. L'évolution d'un réseau dans un état propre est ainsi synchronisée et périodique.

On dispose aujourd'hui d'algorithmes performants pour le calcul de ces valeurs propres dans les algèbres *min-plus* et *max-plus*. Si tout couple de transitions pris au hasard dans le réseau appartient à un cycle, il n'existe qu'une seule valeur propre. Sinon, chacune des composantes du graphe qui vérifie cette propriété possède une valeur propre distincte. Deux ateliers fonctionnant en série sans que la vitesse de traitement du second soit prise en compte dans

la production du premier, par exemple, ont un taux de production découplé et chacun adopte un rythme de fonctionnement périodique correspondant à des valeurs propres différentes. Il faut alors ralentir la production du premier atelier en fonction du rythme du second à l'aide d'un circuit *ad hoc* pour que les deux fonctionnent de concert (le *kanban* repose sur ce principe).

La valeur propre du système détermine le régime périodique adopté à long terme. Le rapport entre les ressources disponibles dans un cycle du graphe d'événements (le nombre de jetons qu'il contient, qui est invariant) et le temps nécessaire pour les traiter (la temporisation de ses places) détermine le taux de production de ce circuit. Le cycle où ce rapport est minimal est nommé circuit critique : il limite en effet la vitesse de fonctionnement du système.

Le circuit critique n'est autre que celui qui minimise le poids des chemins en compteurs (et les maximise en dateurs) dans le calcul des puissances successives de la matrice du système. Au terme d'une période transitoire, les états du système progressent alors périodiquement selon le taux de production du circuit critique. Ce rapport correspond à la valeur propre du système, et donne le rythme naturel sur lequel se cale le fonctionnement du réseau.

L'analyse du comportement dynamique d'un système à événements discrets repose ainsi sur le calcul de la valeur propre de la matrice associée, qui indique le circuit et donc la ressource contraignante. L'identification du « goulot d'étranglement » permet d'optimiser la conception du réseau par l'ajout ciblé de ressources qui améliorent les performances. Ces techniques algébriques s'appliquent aujourd'hui à l'analyse et à la conception de systèmes réels vastes et complexes. La circulation des instructions dans les processeurs est en passe d'être optimisée par ce biais. G. J. Olsder et ses collègues de l'Université de Delft ont étudié de façon semblable le réseau ferroviaire des Pays-Bas. Les protocoles de télécommunication, le trafic urbain et certains ateliers de production font l'objet d'applications de ces techniques. Par ailleurs, même si tous les systèmes synchronisés ne s'expriment pas à l'aide d'équations linéaires dans les algèbres *max-plus* ou *min-plus*, cette approche algébrique facilite la mise au point de simulations performantes.

Les algèbres *min-plus* et *max-plus* ont donc trouvé un champ naturel d'application : les systèmes à événements discrets. De nombreux travaux ont eu lieu sur ce sujet durant la dernière décennie en productique. Par ailleurs, les algèbres *max-plus* et *min-plus* connaissent actuellement de grands développements dans des domaines aussi variés que la théorie des probabilités, la recherche opérationnelle, la physique et les mathématiques pures.

Guy COHEN est chercheur à l'École nationale des Ponts-et-Chaussées. Jean-Pierre QUADRAT et Stéphane GAUBERT sont directeurs de recherche à l'Institut national de recherche en informatique et automatique.

F. BACCCELLI, G. COHEN, G. J. OLSDER ET J.-P. QUADRAT, *Synchronization and linearity. An algebra for discrete event systems*, Wiley & Sons, 1992.

G. COHEN, D. DUBOIS, J.-P. QUADRAT ET M. VIOT, *A linear system-theoretic view of discrete event [...] for performance evaluation in manufacturing*, IEEE Transactions on automatic control, vol. ac-30, n° 30, pp.210-220, 1985.

<http://maxplus.org>