

## Raisonnement causal en physique qualitative

Philippe Dague\* et Louise Travé-Massuyès<sup>◇</sup>

*Résumé* : La Physique Qualitative est une branche de l'Intelligence Artificielle née dans les années soixante-dix, qui s'attache à prédire et expliquer le comportement des mécanismes physiques en termes qualitatifs, donc de manière symbolique. La causalité y intervient comme une théorie de « comment les systèmes fonctionnent » qui fait le lien entre raisonnement sur la structure et raisonnement sur la fonction. Elle joue un rôle dans les tâches de conception, de commande, de diagnostic et d'explication des systèmes physiques. Plusieurs méthodes ont été développées pour inférer des explications causales à partir du modèle d'un système physique. Nous en décrivons essentiellement trois : celle de la causalité mythique de Kleer et Brown, qu'on peut qualifier d'intuitive en ce qu'elle propage de proche en proche le sens d'évolution des variables en réponse à une perturbation ; celle de l'ordonnement causal d'Iwasaki et Simon, qu'on peut qualifier de calculatoire en ce qu'elle assimile enchaînement causal et enchaînement de calcul des valeurs des variables et procède d'une analyse structurelle des équations sans référence au système physique sous-jacent qu'elles modélisent ; celle des graphes de liaisons, issue de l'Automatique, qu'on peut qualifier de méthodologique en ce qu'elle propose un langage unifié à caractère graphique de modélisation des systèmes physiques en s'appuyant sur un formalisme très général en termes de variables d'effort et de flux et de transfert de puissance. Nous illustrons ces méthodes sur des exemples communs afin de les comparer et nous rapportons le vif débat qui a opposé les tenants des deux premières méthodes à propos de la détermination de la causalité en présence de boucles de rétroaction.

*Mots-clés* : raisonnement qualitatif, relation causale, envisionnement, confluence, causalité mythique, heuristique, rétroaction, ordonnancement causal, système complet, causalité intégrale, causalité différentielle, système hybride, graphe de liaisons, effort, flux, puissance.

*Abstract*: **Causal reasoning in qualitative physics.** Qualitative Physics is a field of Artificial Intelligence that emerged in the seventies, which aims at predicting and explaining the behaviour of physical mechanisms in qualitative terms, i.e., in a symbolic way. Causality plays a part as a theory of "how systems work" that links the reasoning about structure and

---

\* LIPN-UMR 7030

<sup>◇</sup> LAAS-CNRS

about function. It plays a role in the tasks of designing, controlling, diagnosing, and explaining physical systems. Several methods have been developed to infer causal explanations from the model of a physical system. We describe mainly three of them: mythical causality of de Kleer and Brown, that can be called intuitive as it propagates step by step the direction of change of the variables in response to a perturbation; causal ordering of Iwasaki and Simon, that can be called computational as it compares a causal sequence with a computation sequence of values of variables and originates in a structural analysis of the equations without reference to the underlying physical system that they model; bond graphs, coming from Control Theory, that can be called methodological as they offer a unified graphical language for modelling physical systems, based on a quite general formalism in terms of effort and flow variables, and energy exchange. We illustrate these methods with common examples in order to compare them and we relate the animated debate that opposed the defenders of the two first methods about assigning causality in the presence of feedback loops.

*Keywords:* qualitative reasoning, causal relation, envisioning, confluence, mythical causality, heuristics, feedback, causal ordering, self-contained system, integral causality, differential causality, hybrid system, bond graph, effort, flow, power.

## 1. INTRODUCTION

La causalité est un des concepts essentiels pour raisonner sur les systèmes physiques. Cet article<sup>1</sup> présente la causalité telle qu'elle est traitée en Physique Qualitative, domaine de l'Intelligence Artificielle, en se focalisant sur l'explication du comportement des systèmes physiques comme nous le détaillerons en section 2. Il est à noter que le raisonnement causal en IA symbolique a donné lieu à de nombreux travaux, à l'interface en particulier des Sciences Cognitives ou de la Linguistique, qui débordent de loin le raisonnement sur les systèmes physiques et dont il ne sera pas question ici.

Le fait de savoir si la causalité est un artefact du raisonnement humain ou recouvre une réalité physique ne sera pas débattu. De façon plus pragmatique, il sera admis que les humains s'attendent fréquemment à avoir des explications en termes causals et que, dans ce qui nous occupe, la causalité joue un rôle central dans le raisonnement sur les systèmes physiques (on s'intéressera essentiellement aux systèmes artificiels conçus et fabriqués par l'homme) et est couramment utilisée par les ingénieurs pour appréhender comment un système fonctionne (ou ne fonctionne plus

---

<sup>1</sup> C'est une version raccourcie et assez largement remaniée du chapitre 7 de l'ouvrage *Modèles et raisonnements qualitatifs*, 2003, des mêmes auteurs.

comme il le devrait) dans des tâches comme la conception ou le diagnostic. Bien que ces tâches de compréhension requièrent en général d'être suivies par des tâches d'intervention (réparation, reconfiguration, accommodation de la commande, etc.), il sera peu question de ces secondes dans cet article et donc la notion d'action (et la théorie des actions) sera peu évoquée. L'article se focalise sur les divers formalismes utilisés pour représenter la causalité et les raisonnements inférentiels qui leur sont attachés.

Dans les domaines de l'ingénieur, il existe souvent des théories fortement structurées qui sous-tendent les comportements des systèmes. C'est dans le sens d'augmenter ces théories avec une connaissance causale supplémentaire que la communauté de Raisonnement Qualitatif a travaillé, en proposant notamment des procédés systématiques pour la découverte des relations causales.

Pour un système physique donné, pour lequel on dispose d'un modèle défini par un ensemble de contraintes liant un ensemble de variables, le problème de l'« ordonnancement causal » se pose comme celui de déduire un ensemble de relations causales (encore appelées *influences*) entre les variables intervenant dans le modèle, ayant spécifié le sous-ensemble de variables exogènes (celles sur lesquelles on peut intervenir – c'est là un des seuls rapports que l'on pourrait établir avec la conception interventionniste de la causalité – ou qui sont fixées par l'environnement du système).  $x R y$  est interprété comme «  $x$  influence  $y$  » ou encore « les événements survenant à  $x$  influencent les événements survenant à  $y$  », sans que les événements ne soient spécifiés.

Les trois principes suivants sont généralement admis comme étant à la base de toute influence causale : ordre temporel, localité et nécessité. Si ces principes, en accord avec l'intuition, fournissent des guides évidents pour ce qui est de la causalité sous-tendant une équation différentielle (description d'un comportement dynamique), ils sont souvent peu utiles dans le cas d'équations algébriques (processus à l'équilibre), en particulier lorsque ceux-ci mettent en jeu des boucles de rétroaction. Il est intéressant de noter que ce dernier point déclencha une polémique entre plusieurs auteurs tenants de méthodes différentes, livrée à la communauté par plusieurs articles successifs de « réponse à » (Iwasaki et Simon, 1986a ; de Kleer et Brown, 1986 ; Iwasaki et Simon, 1986b). Les arguments échangés dans cette polémique sont rapportés en section 5.

Étant donnée la formulation du problème, il est important de noter que l'on se positionne dans un monde clos par le modèle équationnel. Le problème de l'exhaustivité de l'ensemble des relations causales obtenues est donc renvoyé au niveau de la théorie sous-tendant ce modèle.

Cet article décrit et compare les principales méthodes mises en œuvre dans le cadre du raisonnement qualitatif :

- la causalité mythique de de Kleer et Brown (1984 ; 1986) ;
- l'ordonnement causal d'Iwasaki et Simon (1986a ; 1986b ; 1994 ; Iwasaki, 1988) ;
- les graphes de liaison (Top et Akkermans, 1991 ; Dauphin-Tanguy, 2000).

La causalité mythique peut être vue comme une méthode *intuitive* qui cherche à propager de proche en proche le sens d'évolution des variables en réponse à une perturbation. Pour ce faire, elle s'appuie sur des heuristiques. L'ordonnement causal peut être qualifié de méthode *calculatoire* puisque sont confondus enchaînement causal et enchaînement des calculs des valeurs des variables, sans référence particulière au système physique sous-jacent. Enfin, nous qualifierons l'approche par graphes de liaisons d'approche *methodologique pour la physique* en ce qu'elle propose un langage unifié à caractère graphique de modélisation qui s'appuie sur une catégorisation des phénomènes physiques.

Des exemples sont fournis pour illustrer ces méthodes et les comparer sur divers systèmes artificiels et pour différents types de tâches.

## 2. RAISONNEMENT QUALITATIF SUR LES SYSTÈMES PHYSIQUES ET CAUSALITÉ

Le Raisonnement Qualitatif ou Physique Qualitative – domaine de l'Intelligence Artificielle dont l'émergence remonte au début des années soixante-dix – se propose de modéliser le monde physique, caractérisé par des quantités continûment variables dans le temps, de manière symbolique. Il s'agit de prédire et expliquer le comportement d'un mécanisme physique en termes qualitatifs, tout comme le fait un ingénieur utilisant des ordres de grandeur ou un raisonnement sur les tendances. Ceci sous-entend de pouvoir produire une prédiction qualitative du comportement du système physique mais aussi d'être capable de construire une explication causale du pourquoi de ce comportement qualitatif.

Les objectifs originaux prétendaient donc apporter une nouvelle physique, plus simple que la physique classique – le terme Physique Naïve fut utilisé à l'origine – mais capable de retenir les distinctions de comportement significatives d'un point de vue qualitatif sans faire intervenir les mathématiques liées aux équations différentielles et de produire des explications causales de compréhension aisée et intuitives. Les tâches principales que le raisonnement qualitatif aborde sont le diagnostic, la conception, la commande et l'explication de systèmes physiques artificiels. Pour cela, des méthodes telles que l'algèbre symbolique qualitative, l'analyse, la

modélisation, l'«*envisionnement*», la simulation qualitative et le raisonnement causal ont été développées.

Si les objectifs du raisonnement qualitatif débordent de loin ceux du raisonnement causal, la causalité y tient donc une place de choix tout en s'étant avérée de loin le thème le plus controversé. Nous rapporterons en section 5 la polémique à laquelle se livrèrent plusieurs auteurs tenants de méthodes différentes. La causalité intervient comme une théorie de «*comment les systèmes fonctionnent*» qui fait le lien entre raisonnement sur la structure et raisonnement sur la fonction, tous deux à l'œuvre dans la plupart des tâches de raisonnement sur les systèmes physiques artificiels. C'est ainsi que la reconnaissance nécessite d'obtenir une compréhension de la fonction à partir de la structure, la conception de synthétiser une structure qui réalise une certaine fonction et le diagnostic (qui a constitué très tôt un domaine d'application privilégié du raisonnement causal (Brown, 1976 ; Davis, 1983)) d'identifier comment la structure ne réalise plus la fonction donc d'inférer quels changements fonctionnels vont résulter d'éventuels changements structurels. La causalité permet de découvrir plus d'information sur le comportement (par exemple en cas de rétroaction, qui est une propriété du fonctionnement, pas du comportement).

Du point de vue des sciences cognitives, le raisonnement causal est à rapprocher d'un type de raisonnement plus général sur les systèmes physiques appelé «*envisionnement*». C'est de Kleer qui introduit le premier ce terme, à propos d'un système de résolution de problèmes en mécanique classique (de Kleer, 1977), avec la signification d'être capable de prédire grossièrement, à partir d'une connaissance qualitative, le comportement d'un système physique dans une situation donnée, par exemple la chute d'un objet sans support ou le rebond sur le sol d'une balle en caoutchouc. Techniquement, l'envisionnement engendre, sous forme d'un arbre, des successions d'«*instantanés*» décrivant ce qui pourrait arriver. Une vision alternative à celle d'arbre de comportements est celle de graphe de transitions, encore appelé *envisionnement total*, dont les chemins représentent implicitement les comportements qualitatifs possibles. Formellement l'envisionnement total d'un modèle décrit par un ensemble de contraintes qualitatives (appelé usuellement *Equation Différentielle Qualitative* ou EDQ dès qu'y intervient la dérivée par rapport au temps) est constitué de tous les états qualitatifs possibles de l'EDQ et de toutes les transitions possibles entre ces états. La valeur qualitative d'une variable temporelle est ici fournie par la valeur symbolique de la variable dans son espace de définition fini, accompagnée de son sens de variation (signe de sa dérivée), et un état qualitatif est donné par les valeurs qualitatives de toutes les variables. C'est dire que deux états qualitatifs distincts diffèrent par la valeur qualitative d'au moins une variable. Un état qualitatif

possible de l'EDQ est un état qualitatif qui satisfait les contraintes qualitatives définissant cette EDQ. Une transition possible est une transition qui respecte au niveau qualitatif (en limitant les successeurs possibles d'une valeur qualitative) les théorèmes de la valeur intermédiaire et de la valeur moyenne, découlant du caractère continûment différentiable des variables en fonction du temps.

L'envisonnement d'un système modélisé par un ensemble de contraintes qualitatives revêt un caractère ambigu dû à la nature qualitative du modèle, par essence incomplète. Cette ambiguïté se manifeste de deux manières. La première est le branchement d'un état en plusieurs états successeurs possibles dans l'envisonnement. Considérons par exemple un tube en U formé de deux réservoirs A et B reliés en leurs fonds par un tuyau. Partant de l'état initial où A est plein et B vide, on parvient à trois états d'équilibre possibles suivant l'ordre temporel dans lequel surviennent les deux événements : le flux de A vers B atteint 0, le liquide dans B atteint le niveau maximum. Ces états sont : le réservoir B est partiellement rempli, le liquide dans B s'arrête exactement au niveau maximum, le réservoir B déborde. Ici, l'incomplétude du modèle qualitatif (on ne connaît pas les tailles respectives de A et B) ne permet pas de déterminer quel comportement sera le véritable. Mais on peut aisément montrer que chacun des trois comportements arrive effectivement, pour différents tubes en U. Cette propriété n'est hélas pas vraie en général. La seconde source d'ambiguïté est en effet la présence de faux comportements, dans le sens où ils ne sont l'abstraction qualitative d'aucun comportement réel d'un système satisfaisant les contraintes qualitatives du modèle et les conditions initiales.

En résumé, l'envisonnement prédit l'ensemble des comportements possibles du système, cohérents avec la représentation qualitative sous-jacente. Mais, comme on l'a souligné, le raisonnement qualitatif a pour but, non seulement de prédire, mais aussi d'expliquer. Il est tentant de s'appuyer sur l'envisonnement pour la tâche d'explication, par exemple de considérer que la trace d'exécution de tout algorithme de prédiction fournit une explication, ou qu'une explication se ramène à une preuve dans un système de déduction naturelle. Mais on s'aperçoit rapidement (notamment en cas de preuve indirecte nécessitant le recours au raisonnement par l'absurde) que les étapes d'une telle explication ne suivent en général aucune notion d'ordre causal. Suivant de Kleer, les difficultés proviennent notamment du fait que la notion de causalité que nous utilisons pour raisonner est *mythique* (cf. le temps mythique introduit par de Kleer et Brown, 1984) et ressemble bien peu à ce qui se passe vraiment. Ainsi les arguments causaux résultants sont souvent des *rationalisations*, puisque les ambiguïtés sont résolues de manière empirique plutôt qu'analytique.

Il faut donc distinguer ce genre d'explication, qui explique *pourquoi* le système doit se comporter de telle façon, de l'explication causale, qui explique *comment* le système se comporte et qui décrit comment les comportements des constituants contribuent au comportement global du système. Envisonnement et explication causale restent cependant liés dans l'approche de de Kleer (de Kleer et Brown, 1986) qui, pour produire une explication causale, s'appuie sur l'enchaînement des contraintes réalisé lors de la propagation des valeurs des variables pour produire l'arbre d'envisonnement.

De leur côté, Iwasaki et Simon reprennent des travaux effectués dans le domaine de l'économétrie (Simon, 1952 ; 1953) pour aborder le problème de manière totalement différente, sans aucune référence à un quelconque envisonnement ou prédiction qualitative du comportement. Leur but est en effet de proposer une analyse formelle, directement à partir des équations structurelles du système sans se préoccuper de leur obtention, c'est-à-dire de la modélisation, et même sans se référer au dit système, produisant un ordre causal entre les variables qui explique les rationalisations généralement utilisées par l'homme.

Tous considèrent que les modèles des *composants* (selon de Kleer) ou *mécanismes* (selon Iwasaki et Simon) tiennent en général sous une hypothèse quasi-statique, ce qui empêche *a priori* l'expression de la façon dont la causalité rend compte des changements qui interviennent entre deux états d'équilibre du système. Chacun résout ce problème à sa manière, comme il sera détaillé dans les sections 3.4 et 4.1.

Tous admettent cependant que ce problème pourrait être abordé par une *approche physique*, en modélisant les phénomènes physiques à un niveau plus fin de granularité temporelle, ce qui conduirait à expliciter les comportements se déroulant entre deux équilibres. Certains auteurs suivent cette ligne, notamment Rose et Kramer (1991). Leur système QAF (*Qualitative Analysis of Causal Feedback*) accepte en effet comme entrée un modèle du système physique sous forme d'un ensemble d'équations différentielles (qualitatives). Ainsi, lorsque certains composants sont décrits par des équations algébriques, ils forcent la structure du modèle en les remplaçant par des équations différentielles, devant faire appel à des connaissances physiques sur les différentes échelles de temps des phénomènes mis en jeu. Par exemple, l'équation du débit  $Q$  en fonction des pressions à l'entrée  $P_{in}$  et à la sortie  $P_{out}$  pour un fluide incompressible passant au travers d'un filtre ou d'un tuyau est donnée par  $Q = k(P_{in} - P_{out})^{1/2}$ . Elle sera transformée en  $dQ/dt = k(P_{in} - P_{out})^{1/2} - Q$ , où est un petit paramètre représentant la compressibilité du fluide.

Dans cet article, nous ne nous intéresserons pas à l'approche physique, que nous considérons tenir davantage de la physique

classique, pour nous focaliser sur les approches qui se proposent de produire une explication causale en prenant comme base un modèle donné du système, éventuellement quasi-statique.

Nous ne développons pas les approches – nombreuses – faisant usage de modèles causals posés *a priori*. Par exemple, QPT (Qualitative Process Theory) de Forbus (1984 ; Forbus et Gentner, 1983) qui est résolument reconnue comme une approche causale demande, dès la phase de modélisation, à ce que soient répertoriées les influences causales entre les différents processus qui sous-tendent le comportement d'un composant. Mais Forbus ne propose pas de méthode automatique pour cette analyse préliminaire. Et il est vrai que ces modèles causals sont en pratique développés à la main.

### 3. LA CAUSALITÉ MYTHIQUE

#### 3.1 Intuitions sous-jacentes à la causalité mythique

De Kleer et Brown (1986) voient la causalité comme une composante du raisonnement de sens commun (intuitions de la façon dont les systèmes fonctionnent), qui consiste à postuler que *le comportement d'un système résulte d'interactions ordonnées temporellement, de cause à effet, entre composants voisins*.

Ils veulent ainsi définir une notion de causalité qui permette de rendre compte d'un comportement décrit en termes explicatifs (où explication est vue ici comme étapes d'une preuve) de manière causale comme dans un diagramme d'états. À l'image de l'environnement, ils construisent un tel diagramme à partir des états du système et des transitions entre ces états.

Mais la notion d'état est ici différente de celle d'état qualitatif utilisé en environnement et recouvre plutôt celle de mode qualitatif de fonctionnement. Les modèles qualitatifs sont en général définis par morceaux car il est souvent impossible de caractériser fidèlement le comportement d'un composant sur l'ensemble de sa plage de fonctionnement par un unique ensemble d'équations qualitatives. Chaque morceau, que nous appellerons mode opératoire, est spécifié de façon unique par une contrainte où ne figurent que des opérateurs d'égalité ou d'inégalité entre variables (à l'exclusion d'opérateurs temporels comme la dérivée par rapport au temps). Le comportement dans chaque mode opératoire est alors décrit par un ensemble de *confluences*, c'est-à-dire d'équations qualitatives combinant, à l'aide des opérateurs arithmétiques habituels transposés aux signes, les signes des variables et de leurs dérivées. Le modèle complet s'exprime donc par un ensemble {mode opératoire : [spécification], confluences}.

C'est ainsi que le modèle d'une vanne de section  $S$ , variant entre  $0$  et  $S_{MAX}$ , traversée par un fluide de débit  $Q$  avec une différence de



pression  $P$  (dont l'équation analytique pour  $0 < S < S_{MAX}$  est  $Q = C \operatorname{sgn}(P) S \sqrt{2|P|}$  où est la masse volumique du fluide et  $C$  un coefficient) s'écrit avec trois modes opératoires (on note  $[X]$  le signe de  $X$  et  $\partial X$  celui de  $\frac{dX}{dt}$ ) :

- ouverte :  $[S = S_{MAX}], [P] = 0, \partial P = 0$
- en\_fonctionnement :  $[0 < S < S_{MAX}], [P] = [Q], \partial P + [P] \partial S - \partial Q = 0$
- fermée :  $[S = 0], [Q] = 0, \partial Q = 0$

La confluence dans le mode en fonctionnement est dite *mixte* car elle a la forme d'une somme algébrique de produits de variables qualitatives. On peut toujours se ramener à des confluences *pures*, c'est-à-dire sommes algébriques de variables, en ajoutant des modes opératoires. Par exemple, le mode en\_fonctionnement se décompose en trois modes avec des confluences pures :

- en\_fonctionnement\_+ :  $[0 < S < S_{MAX}, P > 0], [P] = [Q], \partial P + \partial S - \partial Q = 0$
- en\_fonctionnement\_0 :  $[0 < S < S_{MAX}, P = 0], [P] = [Q], \partial P - \partial Q = 0$
- en\_fonctionnement\_- :  $[0 < S < S_{MAX}, P < 0], [P] = [Q], \partial P - \partial S - \partial Q = 0$

Un état du système est défini par l'assignation d'un mode opératoire à chacun de ses composants, donc par un ensemble de confluences. Déterminer le comportement intra-état revient à résoudre ces confluences.

Le comportement inter-états est quant à lui décrit par les transitions possibles entre états. Les transitions légitimes sont caractérisées par la possibilité de satisfaire les spécifications d'un état et déterminées à partir du calcul qualitatif. Considérons par exemple un régulateur de pression (la description complète en sera donnée Figure 3.1) dont la vanne est un composant. Supposons que cette vanne soit dans le mode en\_fonctionnement\_+ et que la pression d'entrée du régulateur diminue. La régulation entraîne une augmentation de  $S$  ( $\partial S = +$ ), ce qui offre la possibilité de passer de la spécification  $0 < S < S_{MAX}$  à la spécification  $S = S_{MAX}$ , c'est-à-dire autorise une transition du mode en\_fonctionnement\_+ au mode ouverte. Le processus de construction d'une transition particulière revient à effectuer une intégration à partir des dérivées d'ordre un figurant dans l'état origine (dans l'exemple,  $\partial S$ ) : on retrouvera, sous le nom de causalité intégrale, un concept analogue en ordonnancement causal (voir 4.1.2).

Le diagramme d'états ainsi construit est unique (contrairement aux explications à partir de preuves) et incorpore les propriétés classiques de la causalité (attachée aux transitions, c'est-à-dire inter-états) : *tout changement a une cause et, dans des circonstances identiques, les mêmes causes ont les mêmes effets (nécessité), la*

cause est structurellement proche de l'effet (localité) et la cause précède l'effet (temporalité). Par contre ce diagramme ne fournit pas d'explications causales du comportement *intra-état*. Or ici la résolution des confluences prédit le comportement (comme résultant de la satisfaction des contraintes équationnelles) mais ne dit pas comment le système fonctionne lorsqu'il n'y a pas de changement d'état.

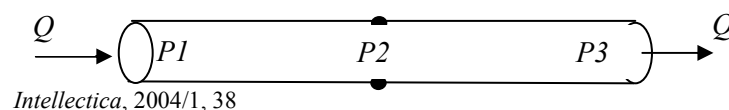
Les auteurs paraphrasent cette situation en disant qu'elle incorpore le principe épistémologique « il y a une raison pour toute chose » aux dépens du principe ontologique « toute chose a une cause ».

Deux problèmes rendent difficile la définition d'une causalité *intra-état* : le temps et l'utilisation du raisonnement par l'absurde (ce dernier est fréquemment utilisé en prédiction lorsqu'on peut réfuter toutes les valeurs qualitatives d'une variable à l'exception d'une, mais il viole le principe de localité car les valeurs ne sont pas obtenues par une propagation de proche en proche mais découlent d'une résolution globale).

Concernant le temps, reprenons l'exemple de la vanne, dont l'équation analytique repose sur l'hypothèse que tout changement de la pression  $P$  s'accompagne d'un changement simultané du débit  $Q$ . Or ceci n'est qu'une approximation. Dans le mode en fonctionnement  $+$ , la confluence reliant le signe des variations des quantités s'écrit :  $\partial P + \partial S - \partial Q = 0$ , qui se simplifie en :  $\partial P - \partial Q = 0$  lorsqu'on suppose  $S$  constant. Ceci traduit le fait que  $P$  et  $Q$  ont le même sens de variation. On peut donc inférer  $\partial Q = +$  de  $\partial P = +$ , mais pas que  $\partial Q = +$  a été causé par  $\partial P = +$  (il n'y a pas de succession temporelle dans l'équation).

L'intuition de la causalité de de Kleer et Brown est que le comportement est produit par des processeurs individuels (un par composant) qui interagissent entre voisins, chacun ne contribuant qu'une seule fois à l'établissement d'un comportement en réponse à une perturbation. Un processeur est conçu pour assurer que la confluence modélisant le composant est satisfaite. Les interactions de transmission de l'information sont interprétées comme des interactions cause-effet. Ces interactions reviennent donc à de la propagation de valeurs de variables entre composants connectés. Évidemment, dès qu'au moins deux variables d'une confluence sont inconnues, ce procédé est inopérant et c'est là qu'intervient le raisonnement par l'absurde.

Considérons l'exemple de deux tuyaux étroits connectés traversés par un même flux :



$$\partial Q = \partial P1 - \partial P2$$

$$\partial Q = \partial P2 - \partial P3$$

Si l'on sait que  $\partial P1 = +$  et  $\partial P3 = 0$ , on ne peut rien déduire directement par propagation. Mais si l'on utilise le raisonnement par l'absurde, on en déduit que  $\partial P2 = +$  car les équations ne peuvent être satisfaites avec  $\partial P2 = -$  ou  $\partial P2 = 0$ .

Pousser l'analyse à des niveaux de détail plus fin, en termes quantitatifs et/ou de modélisation, n'apporte rien. Les équations différentielles quantitatives dont le modèle ci-dessus est une abstraction qualitative nécessitent de la résolution symbolique (système de deux équations à deux inconnues). Descendre à un niveau plus fin de modélisation (faisant intervenir la capacité de stockage des tuyaux et le moment cinétique du fluide) conduit en numérique à une équation différentielle d'ordre 4 avec un comportement oscillatoire amorti, correspondant à des oscillations répétées dans un sens puis dans l'autre, violant le principe d'une seule interaction par processeur. En qualitatif, cela conduit à un diagramme d'états avec 81 états où les ambiguïtés ne peuvent être résolues. D'où l'idée de changer l'interprétation des lois physiques plutôt que le niveau d'analyse pour obtenir la causalité.

### 3.2 Principes de la causalité mythique

Les modèles des composants reposent en général sur l'hypothèse que le système est quasi-statique, ce qui empêche *a priori* l'expression de la façon dont la causalité rend compte des changements qui interviennent entre deux états d'équilibre du système. C'est pour résoudre ce problème que de Kleer et Brown introduisent ce qu'ils appellent causalité et temps mythiques, en termes de négociations sur la transmission d'informations entre processeurs de composants voisins. La *causalité mythique* décrit la trajectoire entre deux états où les modèles quasi-statiques sont valables, sous forme d'« états de non-équilibre ». Il y est associé le concept de *temps mythique*, qui n'impose qu'un *ordre partiel* entre événements, et qui se déroule instantanément en terme de temps conventionnel. La transition entre les deux états se traduit donc par une succession d'instantanés mythiques, auxquels les lois (modèles quasi-statiques des composants) *peuvent être violées*, la durée conventionnelle associée étant *nulle*. Évidemment, ceci ne suffit pas à spécifier de façon non ambiguë ce qui se passe aux instants mythiques et il faut introduire un ensemble de *critères* pour restreindre les options. Leur choix est guidé par le fait qu'on veut qu'ils aident à reconstruire ce que *devrait avoir été le comportement au niveau dynamique si le monde était causal*.

Le premier critère est de présumer que la causalité mythique respecte les mêmes principes que la causalité, excepté en termes temporels, le temps mythique s'écoulant avec une durée nulle.

Le second critère est que, quel que soit le comportement réel dynamique, partant d'un état d'équilibre, on atteint en un nombre fini d'étapes un état d'équilibre. Par un principe de parcimonie, on suppose ce comportement dynamique le plus simple possible.

En résumé, on part d'un état d'équilibre du système (confluences de tous les processeurs composants satisfaites), une perturbation crée un déséquilibre, le système se rééquilibre en temps mythique jusqu'à ce qu'un état d'équilibre soit à nouveau atteint. *L'action causale est précisément le résultat de la violation des confluences aux instants mythiques.* Chaque processeur composant, s'il est associé à une confluence à  $n$  variables, calcule la nouvelle valeur d'équilibre d'une quelconque d'entre elles dès que celles des  $n - 1$  autres sont connues, et l'on impose que chaque variable ne puisse changer de valeur qu'une seule fois, pour une perturbation donnée. L'ensemble des variables avec leurs nouvelles valeurs d'équilibre croît donc de façon monotone en temps mythique et reste topologiquement connexe vis-à-vis de la structure du système : à tout instant mythique, la frontière de processeurs entre anciennes et nouvelles valeurs d'équilibre est bien définie.

Bien entendu, comme on l'a vu dans l'exemple des deux tuyaux, les processeurs (en fait les variables) peuvent se retrouver bloqués faute de propagation possible. Le raisonnement par l'absurde, ou plus généralement l'énumération des valeurs d'une variable, devient nécessaire. On obtient alors en général plusieurs interprétations causales possibles rendant compte du comportement, puisque le choix du processeur, de sa variable bloquée et de la valeur qualitative assignée est arbitraire, sans avoir de méthode pour les distinguer ou les sélectionner. Afin de contourner cet obstacle, de Kleer et Brown introduisent *trois heuristiques* de choix. Il n'y a pas de justification de ces heuristiques en dehors du fait qu'elles fonctionnent dans la plupart des cas, sans que l'on comprenne pourquoi (propriété du monde physique, de la compréhension humaine?). De plus, de l'ambiguïté peut persister dans certains cas et même des résultats incorrects si le retour arrière n'est pas autorisé. Ces insuffisances rendent impossible une prédiction causale du comportement. C'est pourquoi le système ENVISION de de Kleer construit *tous* les comportements qualitatifs possibles et *toutes* les causalités correspondantes qui satisfont les principes de la causalité mythique. Ceci esquivé les insuffisances ci-dessus, mais le prix à payer est que les causalités obtenues n'ont pas la force contraignante d'une explication à partir de preuve.

L'incapacité de « produire causalement » une justification causale du comportement d'un système n'est pas pour les auteurs un défaut

réduisant ni du point de vue de l'IA ni d'un point de vue psychologique. Ils soulignent qu'il n'y a en effet aucune raison, d'un point de vue psychologique, de penser que le type de résolution de problèmes à mettre en œuvre pour construire une telle justification causale soit en lui-même de nature causale, c'est-à-dire n'ait jamais besoin d'avoir accès à une information globale, à revenir sur une décision, etc. Il se peut que la résolution sous-jacente à la construction d'une justification causale n'ait rien à voir avec le problème quasi trivial d'exécuter une prédiction causale, c'est-à-dire de propager des valeurs le long d'un réseau d'influences causales.

### 3.3 Les trois heuristiques

Les trois heuristiques s'appliquent à tout domaine de la physique et s'expriment en termes d'effort et de flux. Les auteurs pensent (mais ne peuvent le prouver) que ces heuristiques sont valables pour une large catégorie de systèmes, soumis à une seule perturbation.

*Heuristique composant : si un composant est soumis à un changement de valeur d'une variable d'effort à l'un de ses terminaux et qu'aucune autre action sur ce composant n'est connue jusqu'à présent, alors le composant répond comme si les actions inconnues étaient négligeables.*

La loi de Kirchhoff sur les tensions est une instance de cette heuristique. Dans le cas d'une vanne, cette heuristique s'écrit : *un changement de la pression (par rapport à une référence absolue) au terminal #1 cause un changement correspondant de la pression entre les terminaux #1 et #2.* Remarquons que la confluence pour la vanne  $\partial P + \partial S - \partial Q = 0$  n'était pas directement utilisable suite à une augmentation de la pression d'entrée  $P_e$  ( $P = P_e - P_s$ ). L'heuristique composant dit par exemple que  $\partial P_e = +$  cause  $\partial P = +$ , ce qui rendra la confluence utilisable. Le même raisonnement s'appliquerait à partir d'une augmentation de la pression de sortie  $P_s$ , avec cette fois un changement de signe opposé.

*Heuristique conduit : si un conduit est soumis à un changement de valeur d'une variable de flux à l'un de ses terminaux, alors la variable d'effort du conduit (par rapport à une référence commune) subit le même changement.*

Cette heuristique est la seule qui relie variables d'effort et de flux dans un conduit. La loi de Kirchhoff sur les courants en est une instance. Appliquée au conduit de sortie de la vanne, elle s'écrit : *un changement du flux volumique au travers de la vanne cause un changement correspondant de la pression (par rapport à la référence) en sortie.* Par exemple,  $\partial Q = +$  cause  $\partial P_s = +$ .

*Heuristique confluence : soit la confluence d'un composant à  $n > 2$  variables ; si  $m < n - 1$  variables sont connues, les propager comme si toutes les variables inconnues à l'exception d'une seule (donc  $n - m - 1$ ) étaient nulles.*

Cette heuristique est une généralisation des deux précédentes. Il y a  $n - m$  façons de l'appliquer. Certaines peuvent être incorrectes, mais une au moins doit être vraie. Appliquée à la vanne, cette heuristique donne : un changement de la pression entre les terminaux #1 et #2 d'une vanne régie par la confluence  $\partial P + \partial S - \partial Q = 0$  cause un changement correspondant du flux volumique. Par exemple,  $\partial P = +$  cause  $\partial Q = +$ . Cet exemple de la vanne est simple car la seconde inférence que pourrait fournir l'heuristique, c'est-à-dire un changement de signe opposé de la surface  $S$ , est impossible, l'aire étant une variable d'entrée uniquement.

Enchaînons les trois heuristiques dans le cas d'une vanne soumise à une perturbation initiale sous forme d'une augmentation de pression à l'entrée ( $\partial P_e = +$ ). Par l'heuristique composant, ceci se traduit par une augmentation de la pression à travers la vanne ( $\partial P = +$ ). Par l'heuristique confluence appliquée à la confluence de la vanne  $\partial P + \partial S - \partial Q = 0$ , il en résulte une augmentation du flux à travers la vanne ( $\partial Q = +$ ). Par l'heuristique conduit, appliquée au conduit de sortie de la vanne (ou encore par la confluence de la charge en sortie :  $\partial P_s - \partial Q = 0$ ), ceci produit une augmentation de pression en sortie ( $\partial P_s = +$ ).

### 3.4 Rétroaction

La causalité mythique permet de déterminer la présence et l'effet d'une rétroaction, sans traitement additionnel. En fait celle-ci est une caractéristique de comment le système fonctionne et n'est donc détectable que par l'analyse causale. Notons qu'on ne parlera de rétroaction que dans le cas où il existe une boucle structurelle, i.e. existant en termes de composants matériels, et pas dans le cas d'une boucle cause-effet comportementale, qui traduit physiquement une sorte de réflexion, comme le phénomène d'une onde réfléchie.

En parallèle avec les premiers travaux de Kleer et Brown (1984) citons le travail de Williams (1984 ; 1990) qui, avec *l'analyse qualitative temporelle*, développe une technique indépendante du domaine pour analyser le comportement qualitatif des systèmes physiques incluant des rétroactions, technique présentant de nombreuses similarités avec la causalité mythique.

#### Exemple 3.1

Illustrons ces idées avec l'exemple du régulateur de pression, dont le mécanisme est illustré par la Figure 3.1 (de Kleer et Brown, 1984 ; Iwasaki et Simon, 1986a).

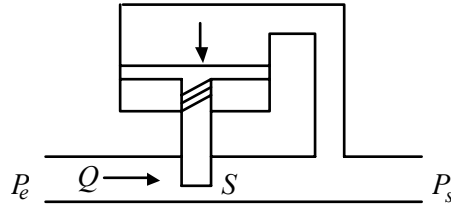


Figure 3.1— Le régulateur de pression

Le comportement de ce système peut être décrit par les équations suivantes :

$$(éq.1) \quad S = - P_s$$

$$(éq.2) \quad P_s = Q$$

$$(éq.3) \quad Q = (P_e - P_s) + S$$

$$(éq.4) \quad P_e = P^*$$

$Q$  est le débit de fluide dans le conduit,  $P_e$  et  $P_s$  sont respectivement la pression à l'entrée et à la sortie du régulateur,  $S$  est la section de l'ouverture réagissant à la pression de sortie. (éq.3) est l'équation de la vanne d'orifice  $S$ ; (éq.2) est l'équation de charge en sortie; (éq.1) est l'équation de régulation de la section par la pression en sortie.  $P_e$  est considérée comme une variable exogène dont la valeur est fixée à  $P^*$  (éq.4).  $\partial P_e$ ,  $\partial P_s$ ,  $\partial Q$  et  $\partial S$  sont des paramètres positifs.

Les confluences associées sont :

$$(éq.5) \quad \partial P_s + \partial S = 0$$

$$(éq.6) \quad \partial P_s - \partial Q = 0$$

$$(éq.7) \quad \partial P_e - \partial P_s + \partial S - \partial Q = 0$$

$$(éq.8) \quad \partial P_e - \partial P^* = 0$$

Considérons la perturbation exogène  $\partial P^* = +$ . Il en résulte par (éq.8) que  $\partial P_e = +$ . À partir de là, aucune propagation n'est possible à travers (éq.5)(éq.6)(éq.7). Seul le raisonnement par l'absurde, en réfutant toutes les autres possibilités, ou la résolution symbolique fournirait la solution unique à ces confluences :  $\partial Q = +$ ,  $\partial P_s = +$ ,  $\partial S = -$ . L'utilisation des heuristiques de propagation permet d'obtenir ce résultat. On a vu en effet au paragraphe précédent, en appliquant les heuristiques sur les équations (éq.7) et (éq.6), qu'une augmentation de pression en entrée cause une augmentation du débit,  $\partial Q = +$ , et une augmentation de pression en sortie, soit  $\partial P_s = +$ . Celle-ci est détectée par le capteur, produisant une diminution de la surface ( $\partial S = -$ ) par la confluence (éq.5). On détecte donc une rétroaction puisque la perturbation a atteint  $\partial S$ , qui était supposé

négligeable lors de l'application de l'heuristique confluence à un antécédent, en l'occurrence à (éq. 7). Si l'on calcule l'effet qu'aurait  $\partial S = -$  si  $\partial P_e - \partial P_s$  était supposé négligeable, on trouve, à partir de la confluence simplifiée  $\partial S - \partial Q = 0$ , que  $\partial Q = -$ , ce qui montre que la rétroaction est négative (bien sûr,  $\partial S = -$  ne peut dominer dans la confluence de la vanne, puisque  $\partial S = -$  ne peut avoir lieu que si  $\partial Q = +$ ). Cette analyse causale fournit donc le graphe suivant :

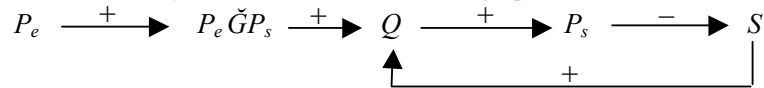


Figure 3.2— Propagation d'une perturbation dans le régulateur de pression par la causalité mythique

#### 4. L'ORDONNANCEMENT CAUSAL

##### 4.1 Systèmes à un seul mode de fonctionnement

Contrairement à de Kleer et Brown, Iwasaki et Simon (1986a ; 1986b ; 1994) déduisent l'ordre causal d'une analyse purement structurelle des équations du système. Leur approche ne nécessite pas la résolution des équations et diffère totalement du processus de recherche de la « causalité mythique » qui détermine le chemin de propagation suivi par une perturbation donnée en entrée. Iwasaki et Simon considèrent que déterminer la structure causale peut être vu comme un problème plus général que celui de déterminer les effets d'une perturbation, qui peut se résoudre par les techniques standard du raisonnement qualitatif une fois que le graphe causal est établi.

L'ordonnancement causal est introduit comme une relation asymétrique au sein des variables et équations d'un système d'équations simultanées représentant le modèle de comportement du système considéré, sans établir un lien avec les constituants physiques sous-jacents (puisque Iwasaki et Simon s'inspirent du domaine économique). Ils parlent plutôt de *mécanismes*, comme des entités conceptuelles définissant les fonctions à partir desquelles le système global doit être expliqué. Les équations correspondant à ces mécanismes sont appelées *équations structurelles* par opposition à des équations qui ne représenteraient pas séparément les mécanismes.

##### 4.1.1. Systèmes statiques

Le système d'équations représentant le comportement d'un système physique doit être « non dégénéré », c'est-à-dire qu'il doit contenir autant d'équations que de variables. L'ordre causal établit un ensemble de relations causales entre les variables. C'est un ordre partiel : il est réflexif, antisymétrique et transitif.



**Définition 4.1**

Un système (qualitatif) statique de  $n$  équations algébriques à  $n$  variables est dit *complet* (*self-contained* en anglais) si tout sous-ensemble de  $k$  équations ( $k \leq n$ ) contient au moins  $k$  variables.

**Définition 4.2**

Étant donné un système complet  $S$ , un sous-système  $s$  de  $S$  qui est aussi complet et qui ne contient pas de sous-ensemble strict complet est appelé *Sous-système Complet Minimal* (SCM). Notons que deux SCMs d'un système complet sont disjoints : aucune variable n'apparaît dans les deux à la fois. Ainsi, chaque SCM peut être résolu indépendamment.

*4.1.1.1. Ordre causal*

Iwasaki et Simon (1986a ; 1986b ; 1994) définissent l'*ordre causal* de la façon suivante : étant donné un système complet  $S$ , considérons  $S_0$  comme l'union de ses SCM, appelés d'ordre zéro. Comme  $S_0$  est complet, les variables de  $S_0$  peuvent être déterminées en résolvant les équations de  $S_0$ . En substituant toutes les occurrences de ces variables dans  $S \setminus S_0$ , on obtient un nouveau système complet, appelé *système dérivé* d'ordre 1. Considérons  $S_1$  l'union de ses SCM, appelés d'ordre 1. La procédure précédente est répétée jusqu'à ce que le dernier système complet dérivé ne contienne aucun sous-système propre complet.

Pour chaque équation  $e_i$  de  $S$ , appelons  $V_i$  l'ensemble des variables apparaissant dans  $e_i$  et  $W_i$  le sous-ensemble de  $V_i$  contenant les variables qui appartiennent aux SCM d'ordre le plus élevé. Alors, les variables de  $W_i$  sont définies comme causalement dépendantes de celles de  $V_i \setminus W_i$ .

La détermination de l'ordre causal est donc équivalente à la recherche des sous-systèmes complets minimaux des structures dérivées successives. Cet ordre causal est unique.

Une boucle de retour correspond à l'existence d'un SCM ayant plus d'une variable. Ainsi, toutes les variables d'un même SCM sont considérées comme mutuellement dépendantes et appartiennent au même niveau causal.

*4.1.1.2. Implémentation à l'aide de la théorie des graphes*

Porté *et al.* (1988) ont montré comment mettre en œuvre les concepts précédents pour la détermination d'un ordre causal dans le cadre de la théorie des graphes. Étant donné un système complet  $S = (E, X)$  constitué d'un ensemble  $E$  de  $n$  équations et d'un ensemble  $X$  de  $n$  variables, le problème de l'ordre causal se ramène à la recherche

d'un couplage parfait<sup>2</sup> dans le graphe biparti associé à  $S$  (Gondran et Minoux, 1979), entre l'ensemble  $E$  d'équations et l'ensemble  $X$  de variables. Il s'agit donc d'associer de manière bi-univoque une variable à une équation.

Le graphe biparti associé à  $S$  est donné par  $G = (E, X, A)$  : chaque équation est représentée par un nœud  $e_i$  et chaque variable par un nœud  $x_j$  et il existe un arc non orienté  $a(j,i)$  entre les nœuds  $x_j$  et  $e_i$  si la variable  $x_j$  apparaît dans l'équation  $e_i$ .

Une fois qu'un couplage parfait  $C$  a été trouvé, chaque équation  $e_i$  peut être interprétée comme un mécanisme qui détermine la valeur de la variable couplée  $x_i$  comme une fonction des autres variables apparaissant dans l'équation.  $x_i$  est alors vue comme *causalement dépendante* des autres variables de l'équation  $e_i$ . Un graphe causal  $G_C = (X, A_C)$  peut être dérivé à partir de  $G$  en orientant les arcs de  $A$  de  $x_i$  vers  $e_j$  si  $a(i,j) \in C$  et de  $e_j$  vers  $x_i$  dans le cas contraire et en agrégeant pour chaque variable couplée le nœud représentant l'équation et celui représentant la variable.

Mais cela n'est pas suffisant pour déterminer l'ordre causal selon Iwasaki et Simon. En effet, le couplage parfait par lui-même ne fournit pas les SCM explicitement.

Porté *et al.* (1988) démontrent que chaque SCM correspond à une Composante Fortement Connexe (CFC) dans le graphe orienté  $G_C$ .

Les résultats précédents montrent que les CFCs de  $G_C$  correspondent aux SCMs de  $S$ . Le graphe causal résultant  $G_{C0}$  est construit à partir de  $G_C$  en agrégeant tous les nœuds à l'intérieur des CFCs de  $G_C$ . Cet ordre causal ne dépend pas du choix du couplage parfait et est unique. Les variables appartenant à la même CFC de  $G_C$ , c'est-à-dire le même SCM de  $S$ , sont mutuellement dépendantes et il n'y a pas d'ordre causal stipulé pour ces variables, i.e. il s'agit d'une boucle.

Par contre, l'ordre causal obtenu directement à partir de  $G_C$  fournit une interprétation de la causalité entre les variables à l'intérieur des CFCs (boucles), qui dépend du couplage parfait utilisé. Dans ce cas, l'ordre causal n'est pas unique, contrairement à celui d'Iwasaki et Simon.

#### Remarque 4.1

Notons que certains auteurs ayant effectué des travaux ultérieurs utilisent l'ordre causal directement obtenu à partir de  $G_C$ . C'est le cas de Travé-Massuyès et Pons (1997) et de Bandekar (1990) qui

---

<sup>2</sup> Etant donné un graphe  $G=(S, A)$ , un *couplage* dans  $G$  est un sous-ensemble d'arêtes  $A' \subset A$  qui sont deux à deux non adjacentes. Un couplage est dit « *parfait* » s'il couvre tous les sommets  $S$  de  $G$ .

préfèrent expliciter les boucles à des fins de diagnostic. C'est le cas également de Lee et Compton (1994).

On peut s'interroger comme le fait Skorstad (1990 ; 1992) sur l'*instabilité contextuelle* de l'ordonnancement causal, c'est-à-dire sa dépendance par rapport aux variables exogènes ou, de manière équivalente, par rapport à l'environnement (ou contexte) dans lequel est placé le système. Mais ceci dépasse la portée de cet article.

#### Exemple 4.1

Reprenons l'exemple du régulateur de pression donné par la Figure 3.1. La matrice d'occurrence  $M$  du modèle donné par les équations (éq.1) à (éq.4) est la suivante :

	$Q$	$P_s$	$S$	$P_e$
(éq.1)		X	X	
(éq.2)	X	X		
(éq.3)	X	X	X	X
(éq.4)				X

X dans une cellule  $m_{i,j}$  signifie que la variable en position  $j$  intervient dans l'équation en position  $i$ .

On constate que (éq.4) forme un SCM d'ordre 0, dont  $P_e$  est la seule variable, et que (éq.1) à (éq.3), où l'on a remplacé  $P_e$  par  $P^*$ , forment un SCM d'ordre 1, dont les variables  $Q$ ,  $P_s$  et  $S$  sont donc mutuellement dépendantes. Pour Iwasaki et Simon, il n'y a pas d'ordonnancement causal des variables  $Q$ ,  $P_s$  et  $S$ :  $P_s$  et  $S$  influencent  $Q$ ;  $P_s$  contrôle  $S$ ; et  $Q$  influence aussi  $P_s$ . La structure causale est donc donnée par la Figure 4.1 :

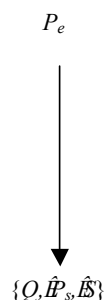


Figure 4.1 — Structure causale du régulateur de pression suivant Iwasaki et Simon (1986a)

#### 4.1.1.3 Utilisation de la statique comparative

On constate que cet ordre causal seul ne permet pas de déterminer le comportement du système suite à une perturbation, telle qu'un accroissement de la pression en entrée, contrairement à la causalité mythique (Figure 3.2). Le recours à la méthode de la statique comparative (« comparative statics ») permet à Iwasaki et Simon (1986a) d'obtenir ce résultat. Il s'agit d'introduire explicitement la dynamique des processus. On peut ici, de façon réaliste, supposer que l'ajustement de l'ouverture de la vanne d'une part et celui du débit d'autre part prennent du temps, alors que l'effet du débit sur la pression de sortie est rapide comparé à ces deux processus et peut être considéré comme instantané. Ceci revient donc à remplacer (éq.1) et (éq.3) par des équations différentielles et à garder (éq.2) inchangée. On obtient, en première approximation (en supposant que les retours de  $S$  et  $Q$  à leurs valeurs d'équilibre suite à une perturbation se font à des vitesses proportionnelles à leurs déviations par rapport à ces valeurs) le système suivant (où  $c_1$  et  $c_2$  sont positifs) :

$$(éq.9) \quad \frac{dS}{dt} = c_1 (P_s - S)$$

$$(éq.10) \quad P_s = Q$$

$$(éq.11) \quad \frac{dQ}{dt} = c_2 (P_e - P_s) + S - Q$$

$$(éq.12) \quad P_e = P^*$$

La statique comparative détermine la réponse nette à une perturbation entre deux états d'équilibre. Il faut donc au préalable s'assurer de la stabilité dynamique du système en analysant les équations ci-dessus. On peut alors perturber le système par  $\partial P^* = +$  et en résolvant symboliquement le système de confluences (éq.5)-

(*éq.8*), on peut déduire que  $\partial P_s = +$  et  $\partial S = -$ . Contrairement à la causalité mythique, qui se contente de propagation, mais a besoin d'heuristiques pour la guider, c'est ici de la résolution symbolique qui doit être appliquée. Enfin, si on obtient les mêmes résultats pour les déviations lors du retour à l'équilibre, il n'y a pas de chemin causal comme dans la Figure 3.2. Voyons à présent comment Iwasaki et Simon ont généralisé l'ordonnancement causal aux systèmes dynamiques.

**4.1.2 Systèmes dynamiques**

Les systèmes dynamiques sont modélisés par des équations différentielles. On peut supposer, sans perte de généralité, qu'il s'agit d'équations du premier ordre, puisque toute équation différentielle d'ordre quelconque se réécrit en un système d'équations différentielles du premier ordre.

**Définition 4.3**

Un système (qualitatif) dynamique de  $n$  équations différentielles du premier ordre à  $n$  variables est dit *complet* si tout sous-ensemble de  $k$  équations ( $k \leq n$ ) contient au moins les dérivées premières de  $k$  variables.

Iwasaki et Simon distinguent deux types de relations causales dans les systèmes dynamiques : la causalité intégrale et la causalité différentielle. La causalité intégrale signifie que chaque variable dépend de sa dérivée :  $t, x(t) \quad x(t \ dt) \quad \frac{dx}{dt} dt$ . C'est ainsi que

(*éq.9*) et (*éq.11*) fournissent les causalités intégrales suivantes :

$$\frac{dQ}{dt} \xrightarrow[\text{I}]{+} Q \qquad \frac{dS}{dt} \xrightarrow[\text{I}]{+} S$$

Figure 4.2— Causalité intégrale (le I signifie « intégrale »)

La causalité différentielle suppose que le système a préalablement été mis sous *forme canonique* équivalente (ce qui est possible par une manipulation algébrique des équations), où une équation différentielle est dite sous forme canonique s'il y a une seule dérivée dans l'équation et que cette dérivée est le seul terme apparaissant dans le membre de gauche de l'équation :  $\frac{dx_i}{dt} = f(x_1, \dots, x_i, \dots, x_n)$ .

Toute équation différentielle sous forme canonique peut alors être interprétée comme un mécanisme déterminant la valeur de la dérivée comme une fonction des variables apparaissant dans le membre de droite. C'est cette interprétation causale naturelle exprimant la

dépendance de chaque dérivée par rapport à toutes les variables apparaissant dans son expression qui est appelée *causalité différentielle*. C'est ainsi que (éq.9) et (éq.11) fournissent les causalités différentielles suivantes :

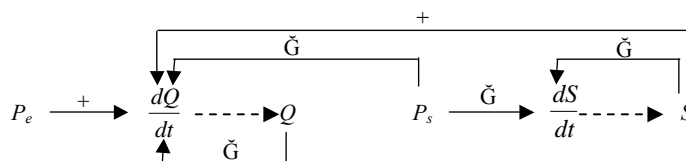


Figure 4.3 — Causalité différentielle (traits pleins) et intégrale (traits pointillés)

L'ordre causal d'un système dynamique complet sous la forme canonique est donné, pour chaque équation, par :

- le lien intégral liant une dérivée et sa primitive ;
- les liens différentiels entre une dérivée et les autres variables de l'équation.

La superposition des liens de causalité différentielle et intégrale illustrée en Figure 4.3 fournit quasiment l'ordonnancement causal pour le régulateur de pression. Seule (éq.10), reliant Q et Ps, n'a pas été traitée, car cette équation est statique. Il faut donc étendre cette technique au cas des systèmes mixtes.

### 4.1.3 Systèmes mixtes

Les structures mixtes sont composées d'équations différentielles et d'équations algébriques. C'est le type de système le plus fréquemment rencontré ; les modèles mixtes sont souvent utilisés à la place de modèles dynamiques purs quand la dynamique de certaines variables peut être considérée comme instantanée par rapport aux autres.

#### 4.1.3.1 Instanciation des systèmes mixtes

Étant donnée l'équation suivante :  $t, x(t) \quad x(t - dt) \quad \frac{dx}{dt} dt,$

Iwasaki et Simon considèrent  $x(t - dt)$  comme une variable exogène pour l'instanciation du système à l'instant  $t$ . L'ordre causal pour un système  $S$  à l'instant  $t$  est obtenu à partir d'un nouveau système

$Inst(S)$  incluant toutes les équations de  $S$  et, pour chaque dérivée  $\frac{dx_i}{dt}$  apparaissant dans  $S$ , une équation constante  $x_i = c$  pour représenter le fait que  $x_i$  est une variable exogène dans  $Inst(S)$ .

**Définition 4.4**

Un système mixte  $S$  de  $n$  équations à  $n$  variables est *complet* si et seulement si :

$S$  est composé d'équations différentielles du premier ordre et d'équations algébriques ;

$Inst(S)$  est un système statique complet lorsque les dérivées des variables sont considérées comme de nouvelles variables distinctes.

Notons que, si  $S$  est statique, on retrouve la définition précédente, mais que si  $S$  est purement dynamique, cette nouvelle définition est plus générale que la précédente : en particulier, elle n'impose pas la mise sous forme canonique et donc pas non plus la causalité différentielle.

**4.1.3.2 Ordre causal**

L'ordre causal d'un système mixte est déterminé par :

L'application des règles d'ordonnement des systèmes statiques à  $Inst(S)$  ;

L'ajout des liens intégrateurs entre chaque dérivée et sa primitive (causalité intégrale).

Illustrons ceci sur le régulateur de pression, en partant des équations (éq.9)-(éq.12), qui forment un système mixte. Son instantiation est obtenue en y ajoutant les deux équations suivantes :

$$(éq.13) \quad S = s$$

$$(éq.14) \quad Q = q$$

où  $s$  et  $q$  sont des constantes. On obtient donc un système statique de 6 équations à 6 variables ( $\frac{dS}{dt}$  et  $\frac{dQ}{dt}$  étant des variables indépendantes), dont il est facile de voir qu'il est complet.

Le système mixte est donc complet. Les équations (éq.12)-(éq.14) forment 3 SCM d'ordre 0, mettant en jeu les variables  $S$ ,  $Q$  et  $P_e$ . L'équation (éq.10), où  $Q$  a été remplacé par  $q$ , forme un SCM d'ordre 1, mettant en jeu la variable  $P_s$ . Enfin, (éq.9) et (éq.11), dans lesquelles  $S$ ,  $Q$ ,  $P_e$  et  $P_s$  ont été respectivement remplacées par  $s$ ,  $q$ ,  $P^*$  et  $q$  forment 2 SCM d'ordre 2, mettant en jeu les variables  $\frac{dS}{dt}$  et  $\frac{dQ}{dt}$ . L'ordonnement causal de ce système statique instancié,

auquel on ajoute les deux causalités intégrales entre  $\frac{dS}{dt}$  et  $S$  et entre  $\frac{dQ}{dt}$  et  $Q$ , fournit l'ordre causal suivant :

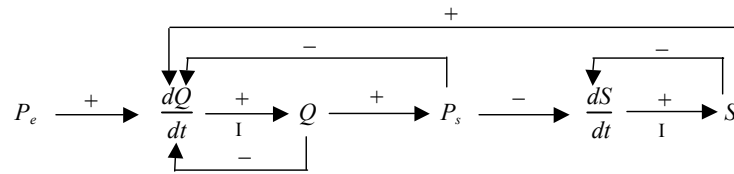


Figure 4.4 — Ordonnement causal du modèle mixte du régulateur de pression

On remarque que, si l'on ne considère pas les 2 variables dérivées, on retrouve comme approximation de cet ordonnancement causal la structure causale obtenue par la causalité mythique à l'aide d'heuristiques (Figure 3.2).

#### 4.2 Systèmes à plusieurs modes de fonctionnement ou hybrides

La majorité des systèmes ont *plusieurs modes opératoires*, ce qui leur procure un caractère *hybride* (Tomlin et Greenstreet, 2002). En effet, ils sont le siège de processus continus, mais incluent cependant des commutateurs (électriques, hydrauliques, etc.) ou sont bouclés sur des contrôleurs de nature discrète. Ils peuvent aussi avoir un comportement hautement non linéaire qu'il est plus aisé de représenter en associant différents comportements à des plages de fonctionnement distinctes. Le modèle équationnel d'un tel système est *modal* dans le sens où des *conditions logiques* sont associées à ses équations. Ces conditions se retrouvent associées aux influences de la structure causale, leur validité spécifiant l'état actif ou inactif des influences.

Travé-Massuyès et Pons (1997) proposèrent une extension de la méthode d'ordonnement causal d'Iwasaki et Simon pour de tels systèmes. Lors d'un changement de mode, la structure causale correspondant au nouveau mode est générée de manière incrémentale, c'est-à-dire en partant de celle dans le mode précédent plutôt que d'être recalculée entièrement. On parle d'ordre causal incrémental.

##### Définition 4.5 (Cohérence d'un mode opératoire)

Un mode opératoire est dit *cohérent* si et seulement s'il est logiquement et physiquement cohérent, c'est-à-dire que l'ensemble des conditions logiques définissant ce mode est satisfiable et qu'elles sont toutes simultanément réalisables (d'un point de vue physique).



**Définition 4.6** (Complétude)

Un système  $S$  à modes multiples est *complet* si et seulement si ses parties statiques et dynamiques sont complètes pour tous les modes opératoires.

Le nombre de modes opératoires différents du système est obtenu en multipliant le nombre de modes différents pour chaque composant physique. On doit cependant noter que parmi ces modes, certains peuvent être incohérents (voir définition 4.5). On définit un mode cohérent comme une *configuration*.

Considérons le système  $S$  dans une configuration donnée — prise comme la configuration initiale sans perte de généralité — dont le modèle est constitué par un ensemble d'équations et de variables. Pour cette configuration, la méthode d'Iwasaki et Simon (1993) permet de générer un graphe causal initial  $G_{C0}$ . Un changement de configuration est défini par le changement des valeurs de vérité d'un ensemble de conditions. On définit les *macro-conditions*  $C$  et  $C^*$  comme l'union des conditions dont la valeur de vérité passe de 1 à 0, et de 0 à 1 respectivement. Ceci revient à remplacer les équations conditionnées par les conditions de  $C$  par celles soumises aux conditions de  $C^*$ . Alors, les résultats de Travé-Massuyès et Pons (1997) permettent de déterminer le sous-graphe causal minimal de  $G_{C0}$  devant être réévalué (i.e. non applicable à la configuration courante) et le sous-graphe venant le remplacer.

**Exemple 4.2**

Reprenons l'exemple du régulateur de pression intégré dans un système d'alimentation. Le régulateur alimente un réservoir ayant, suivant le mode, une ou deux extractions de fluide, comme illustré sur la Figure 4.5. La première est constamment ouverte alors que la seconde inclut un composant (hachuré sur la figure) procurant un caractère hybride au système global. Les débits dans les deux extractions 1 et 2 sont notés  $Q_1$  et  $Q_2$  et leurs pressions de sortie  $P_{s1}$  et  $P_{s2}$  sont considérées constantes, égales à la pression atmosphérique  $P_{atm}$ .

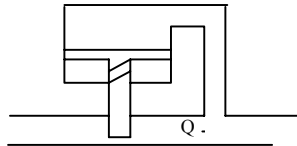


Figure 4.5 — Le régulateur de pression intégré à un système d'alimentation

Le niveau de fluide dans le réservoir est noté  $H$  et la pression en son fond  $P_R$ . Le composant hachuré est une vanne tout ou rien à seuil. Elle est fermée tant que la différence de pression à ses bornes ( $P_R - P_{s2}'$ ) reste inférieure à un seuil au-delà duquel elle s'ouvre. On définit la condition  $C$  comme étant vraie lorsque  $P_R - P_{s2}' < \dots$ , fausse dans le cas contraire et l'on note  $C^* = \neg C$ .

Les équations de ce système sont données par l'ensemble des équations (éq.9)-(éq.14) du régulateur (modèle mixte) déjà utilisées auxquelles on ajoute les équations (éq.15), avec son équation quasi-exogène (éq.16), et (éq.17) pour le réservoir, une équation pour chaque mode de comportement de la vanne à seuil (éq.18) (conditionnée par  $C$ ) et (éq.19) (conditionnée par  $C^*$ ), une équation (éq.20) pour la 1<sup>ère</sup> extraction et une équation (éq.21) pour la 2<sup>ème</sup> extraction (conditionnée par  $C^*$  car cette branche est fermée dans le cas contraire) ainsi que deux équations exogènes (éq.22) et (éq.23).  $k_d$ ,  $\dots$  et  $\dots$  sont des coefficients constants qui dépendent des caractéristiques du fluide, du réservoir, de la vanne à seuil et des extractions.

Réservoir :

$$(éq.15) \quad \frac{dH}{dt} = Q - Q_1 - Q_2$$

$$(éq.16) \quad H = h$$

$$(éq.17) \quad P_R = \rho H$$

Vanne à seuil :

$$(éq.18) \quad Q_2 = 0 \quad \text{sous } C$$

$$(éq.19) \quad P_R - P_{s2}' = k_d \quad \text{sous } C^*$$

1<sup>ère</sup> extraction :

(*éq.20*)  $Q_1 = {}_1(P_R - P_{s1})$

2<sup>ème</sup> extraction :

(*éq.21*)  $Q_2 = {}_2(P_{s2}' - P_{s2})$  sous C\*

(*éq.22*)  $P_{s1} = P_{atm}$

(*éq.23*)  $P_{s2} = P_{atm}$  sous C\*

Considérons la configuration initiale comme correspondant à la condition C. Dans cette configuration, les équations (*éq.19*), (*éq.21*) et (*éq.23*) ne font pas partie du modèle. La structure causale obtenue, donnée par la Figure 4.6, inclut celle du régulateur seul (Figure 4.4) qui influence de manière positive le réservoir.

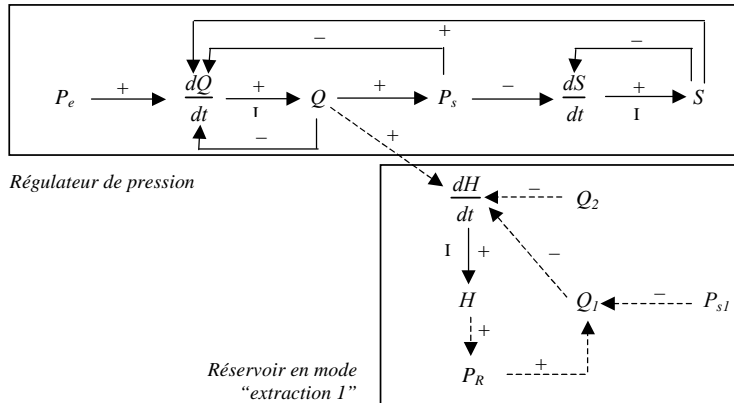


Figure 4.6 — Structure causale pour la configuration initiale (condition C)

Passons à la nouvelle configuration correspondant à la condition C\*. Les équations (*éq.19*), (*éq.21*) et (*éq.23*) viennent remplacer l'équation (*éq.18*). Les résultats de Travé-Massuyès et Pons (1997) permettent de déterminer le sous-graphe qui doit être réévalué, indiqué par les arcs en pointillés sur la Figure 4.6, ainsi que le sous-graphe qui vient le remplacer dans la nouvelle configuration. Il s'agit des arcs en pointillés sur la Figure 4.7. La structure causale concernant le régulateur de pression n'est pas concernée par le changement de configuration et seule la partie concernant le réservoir et les extractions est remise en cause.

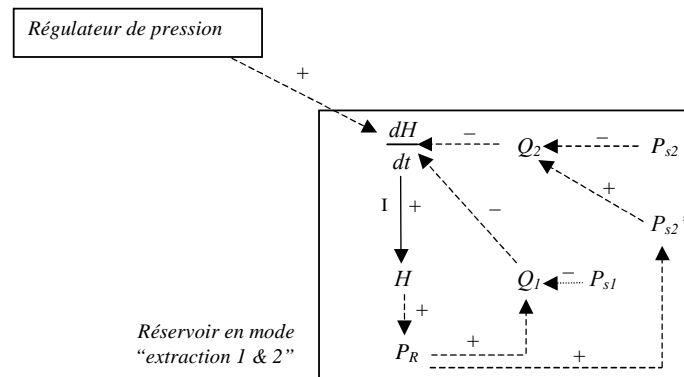


Figure 4.7— Structure causale pour la configuration finale (condition C\*)

### 5. LES BOUCLES DE RÉTROACTION : CAUSALITÉ FICTIVE OU RÉALITÉ ?

La plupart des systèmes physiques incluent des boucles de rétroaction (*feedback loops* en anglais). Lorsque de tels systèmes sont modélisés à un niveau d'abstraction élevé, ces boucles correspondent à des boucles algébriques, c'est-à-dire un ensemble de  $n$  équations algébriques à  $n$  variables inconnues. Les chercheurs s'accordent pour dire que la signification de la causalité n'est pas bien claire dans ce cas. Certains, qui refusent d'abandonner la propriété de nécessité, déclarent, comme Iwasaki et Simon (1994), qu'elle n'a alors aucun sens. D'autres, comme de Kleer et Brown (1984 ; 1986) et Williams (1990), relâchent la propriété de nécessité et disent qu'il peut alors exister plusieurs interprétations causales. Cependant, ces derniers constatent que les ingénieurs ont une interprétation causale préférée qui peut s'appréhender par des heuristiques. Par exemple, sauf argument contraire, on peut préférer que le changement de pression cause le changement de débit plutôt que le contraire (les heuristiques KCL et KVL de de Kleer, 1979). Williams (1990) propose quant à lui une théorie pour déterminer la direction dominante du *flux causal* dans les boucles en exploitant les propriétés mathématiques des éléments capacitifs.

Le problème posé par les boucles provient du fait qu'une explication causale décrit le comportement d'un système comme une séquence d'effets locaux, alors que l'analyse des boucles doit se faire suivant une perspective globale. Ainsi ce problème fut l'objet d'un débat entre Iwasaki et Simon (1986a ; 1986b) d'un côté, et de Kleer et Brown (1984 ; 1986) de l'autre, livré à la communauté par plusieurs articles successifs dans *AI Journal* en 1986.

Les sections 3.4 et 4.1 ont présenté les deux approches concurrentes. Dans l'approche de causalité mythique, la recherche de la causalité sous-jacente s'effectue de manière dynamique lors de la propagation d'une perturbation dans le jeu de contraintes. Au contraire, l'approche de l'ordonnement causal s'effectue hors-ligne en analysant la structure des équations et sans nécessiter des signaux d'entrée.

Il est important de remarquer que, partant du même jeu d'équations, les deux approches établissent le même ordre causal lorsque le système ne contient pas de boucles. Dans ce cas, la propagation « physique » d'une perturbation d'une variable exogène trouve sa correspondance dans les substitutions « calculatoires » de variables induites par l'analyse structurelle des équations. Quand une boucle existe, l'envisonnement détermine toutes les interprétations et la *causalité mythique* sélectionne, grâce à des heuristiques, une interprétation causale préférentielle le long de la boucle, alors que *l'ordonnement causal* s'interdit de donner une telle interprétation causale, tout en proposant de faire une analyse complémentaire inspirée de la méthode de statique comparative des économistes pour déterminer l'effet d'une perturbation (cf. section 4.1.1.3).

Dans leur réponse, de Kleer et Brown (1986) expliquent la source des différences des deux approches dans les différences entre les domaines techniques et le domaine économique et le fait qu'ils rattachent la causalité directement à la structure physique du système, alors qu'Iwasaki et Simon la rattachent à la structure des équations qui décrivent le système. Ils développent plusieurs conséquences de ce fait, en particulier en ce qui concerne leur capacité à produire une interprétation causale pour les boucles. Dans leur réponse, Iwasaki et Simon (1986b) rejettent cet argument, affirmant qu'ils partent des *équations structurelles*, chacune représentant bien un *mécanisme* distinct, les mécanismes étant définis comme des entités conceptuelles définissant les fonctions à partir desquelles le système global doit être expliqué. Il n'y a donc aucune différence entre les deux approches suivant ce point de vue. Ils appuient cet argument en remarquant que de Kleer et Brown utilisent aussi bien des équations « définitionnelles » dans leurs modèles, équations qui ne se rapportent donc pas à des composants ou connexions.

Ainsi Iwasaki et Simon réfutent la plupart des arguments de de Kleer et Brown conséquence de cette (fausse) différence. A part pour le cas des boucles, qu'ils considèrent sans objet, ils s'accordent pour dire que les résultats produits sont identiques mais considèrent que la méthode de causalité mythique est moins générale et que la méthode plus formelle d'ordonnement causal explique pourquoi cette première marche.

Avec le recul nécessaire, ce débat semble assez vain, puisque bien trop fondé sur les intuitions respectives sur la causalité des auteurs en présence. Comme le fait remarquer l'introduction de la partie sur la causalité des *Readings in QR* (Weld et de Kleer, 1990) on peut lui trouver néanmoins une utilité dans le fait qu'il met en évidence la complexité du concept de causalité. Les différences des deux approches ne reposent vraisemblablement pas tant sur le fait d'être un ingénieur ou un économiste, mais sur la classe de tâches pour laquelle la causalité est utilisée. Assigner une causalité dans les boucles pour des tâches typiques d'économie peut être sans intérêt. Assigner une causalité dans les boucles est au contraire de première importance en ingénierie car la plupart des tâches nécessitent de modifier d'une certaine façon la boucle. Que les notions de causalité qu'ingénieurs et économistes utilisent pour réaliser leurs tâches respectives soient différentes ou non est une question d'ordre pragmatique et pas nécessairement d'un intérêt central à la physique qualitative.

## **6. L'APPROCHE PAR GRAPHES DE LIAISONS (« BOND GRAPHS »)**

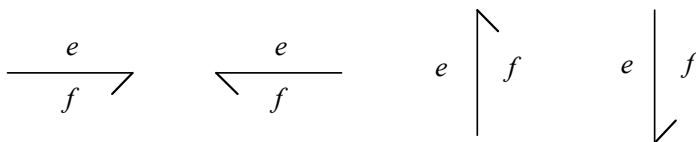
Les graphes de liaisons (Dauphin-Tanguy, 2000) sont proposés pour représenter les systèmes physiques comme un intermédiaire entre la représentation physique et le modèle mathématique. Imaginés par Paynter (1961), formalisés par Karnopp et Rosenberg (1975 ; 1983), arrivés en Europe à la fin des années 1970 par les Pays-Bas et la France, ils sont utilisés depuis longtemps en automatique systémique et sont apparus plus récemment en IA. Ils s'appuient sur la représentation des transferts de puissance au sein du système et sur un formalisme très général de modélisation en termes de variables d'effort et de flux (la puissance étant le produit d'un effort par un flux). La causalité y est représentée localement par l'impossibilité d'imposer simultanément un effort et un flux. C'est un modèle graphique à partir duquel on va pouvoir d'une part, engendrer de manière systématique les modèles mathématiques « classiques » qui sous-tendent le comportement du système physique et d'autre part, exhiber les relations de cause à effet existantes.

C'est souvent pour ce dernier point que certains chercheurs de la communauté du Raisonnement Qualitatif se sont intéressés à cette méthode de représentation (Top et Akkermans, 1991 ; Mosterman, 1997 ; Mosterman et Biswas, 1997 ; Mosterman *et al.* 2000).

### **6.1 Les principes de base**

Les variables généralisées utilisées permettent d'assurer une analogie entre les différents principaux domaines de la physique. Il s'agit des variables d'effort  $e$  et de flux  $f$ , qui sont des variables de

puissance, et de leurs intégrales par rapport au temps, les variables de moment  $p$  et de déplacement  $q$  respectivement, qui sont des variables d'énergie. Le transfert de puissance entre deux sous-systèmes est représenté par une demi-flèche qui correspond au « bond » (liaison) du graphe. Le sens de la demi-flèche correspond au sens positif de la puissance. Ce lien porte les variables d'effort et de flux, dont le produit égale la puissance  $P = e.f$  (le flux est représenté du côté de la demi-flèche).



Le tableau 6.1 indique la signification de ces variables généralisées pour les principaux domaines de la physique.

Domaine	Effort $e$	Flux $f$	Moment $p$	Déplacement $q$
mécanique translation rotation	force $F$ couple	vitesse $v$ vitesse angulaire	impulsion $p$ impulsion angulaire $h$	élongation $x$ angle
électrique	tension $u$	courant $i$	flux magnétique	charge $q$
hydraulique	pression $P$	débit volumique $Q$	impulsion pression	volume $V$
thermique	température	flux d'entropie		entropie
magnétique	force magnéto-motrice	dérivée flux magnétique		flux magnétique
chimique	potentiel chimique	flux molaire		masse molaire

Tableau 6.1—Variables généralisées pour différents domaines de la physique

Les composants physiques sont catégorisés en éléments passifs (au nombre de trois), éléments actifs (au nombre de deux) et éléments de jonction (au nombre de quatre).

Les trois éléments passifs 1-port (qui reçoivent la puissance par un lien unique ; il est parfois nécessaire d'utiliser des éléments passifs multiports) sont : l'élément résistif R, l'élément capacitif C et l'élément inductif I. Ils transforment la puissance fournie en énergie dissipée sous forme de chaleur (R) ou stockée (C et I).

L'élément R modélise tout phénomène physique liant les variables d'effort et de flux.

$$\frac{e}{J} \nearrow R$$

La loi générique  $R(e, f) = 0$  est linéaire dans les cas de la loi d'Ohm en électricité ( $u = R_i i$ ) ou de la loi de frottement en mécanique ( $F = bv$ ) par exemple.

L'élément C modélise tout phénomène physique liant les variables d'effort et de déplacement.

$$\frac{e}{f = dq/dt} \nearrow C$$

La loi générique  $C(e, q) = 0$  est linéaire dans les cas de la loi d'un condensateur en électricité ( $u = q/C_I$ ) ou de la loi d'un ressort en mécanique ( $F = kx$ ) par exemple.

L'élément I modélise tout phénomène physique liant les variables de flux et de moment.

$$\frac{e = dp/dt}{f} \nearrow I$$

La loi générique  $I(p, f) = 0$  est linéaire dans les cas de la loi d'une bobine en électricité ( $\dot{u} = Li$ ) ou de la loi d'une masse inertielle en mécanique ( $p = Mv$ ) par exemple.

Les deux éléments actifs, qui fournissent de la puissance au système, sont les sources d'effort  $Se$  (par exemple, un générateur de tension électrique) et de flux  $Sf$  (par exemple, un générateur de débit volumique hydraulique) qui imposent respectivement les variables d'effort et de flux, indépendamment de la variable complémentaire.

$$Se \longrightarrow \qquad \qquad Sf \longrightarrow$$

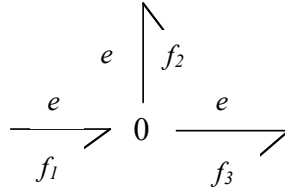


Les quatre éléments de jonction, 0, 1, TF, GY, servent à coupler les éléments passifs et actifs et sont conservatifs de puissance.

La jonction 0 sert à coupler des éléments soumis au même effort (égalité des efforts, nullité de la somme algébrique des puissances, donc des flux). C'est une généralisation à tout domaine physique de la loi de Kirchhoff sur les courants en électricité.

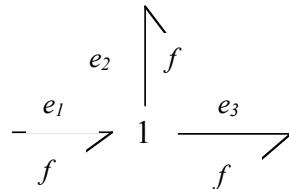
$$e_1 = e_2 = \dots = e_n$$

$$\sum a_i f_i = 0 \quad (a_i = +1 \text{ si la demi-flèche est dirigée vers la jonction et } -1 \text{ sinon}).$$

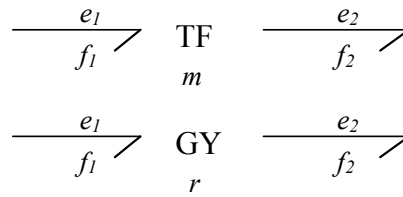


La jonction 1 sert à coupler des éléments parcourus par le même flux (égalité des flux, nullité de la somme algébrique des puissances, donc des efforts). C'est une généralisation à tout domaine physique de la loi de Kirchhoff sur les tensions en électricité.

$$f_1 = f_2 = \dots = f_n \quad \sum a_i e_i = 0$$



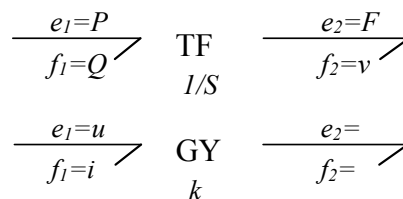
Le transformateur TF et le gyrateur GY sont des éléments 2-port modélisant les phénomènes physiques qui transforment la puissance, tels les transformateurs électriques et les gyroscopes respectivement.



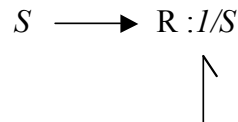
Les relations qui caractérisent TF et définissent  $m$  sont :  $e_1 = me_2, f_2 = mf_1$ .

Les relations qui caractérisent GY et définissent  $r$  sont :  $e_1 = rf_2, e_2 = rf_1$ .

TF et GY sont en particulier utilisés pour modéliser les changements de domaines physiques qui se font sans perte de puissance, comme le couplage hydraulique-mécanique dans un vérin hydraulique ( $S$  étant l'aire du piston) et le couplage électrique-mécanique dans un moteur à courant continu.



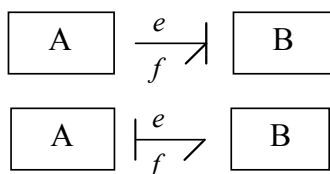
Il existe enfin des liens à puissance nulle, qui correspondent à une liaison transférant de l'information et non pas de la puissance, symbolisée par une flèche entière. Elle peut servir par exemple à représenter des coefficients modulables. C'est le cas dans le régulateur de pression où la valeur de la résistance hydraulique dépend de la valeur de la section de l'ouverture, c'est-à-dire de la valeur du déplacement du piston.



## 6.2 Expression de la causalité

Un modèle par graphe de liaisons permet de faire apparaître explicitement les relations de cause à effet. Lorsque deux sous-systèmes A et B sont couplés et échangent de la puissance, deux

situations sont en effet possibles : A applique un effort  $e$  à B, qui réagit en envoyant à A un flux  $f$  fonction de  $e$  ; A envoie un flux  $f$  à B, qui réagit par un effort  $e$  fonction de  $f$ . Pour représenter graphiquement ces relations, un « trait causal » est placé perpendiculairement à chaque liaison, à l'extrémité voisine de l'élément auquel l'effort est imposé.



L'affectation de la causalité est soumise à des règles regroupées dans Tableau 6.2 et Tableau 6.3.

Causalité	Élément	Loi caractéristique (en général)	Cas linéaire
obligatoire	Se $\longrightarrow \nearrow$	$e$ imposé par Se	
	Sf $\longleftarrow \dashv$	$f$ imposé par Sf	
intégrale	$\longleftarrow \dashv$ C	$e_C = \int_C f_C dt$	$e_C = q_C / C$
	$\longrightarrow \nearrow$ I	$f_I = \int_I e_I dt$	$f_I = p_I / I$
dérivée	$\longrightarrow \nearrow$ C	$f_C = \frac{d}{dt} ( \int_C^{-1} (e_C) )$	$q_C = C e_C$
	$\longleftarrow \dashv$ I	$e_I = \frac{d}{dt} ( \int_I^{-1} (f_I) )$	$p_I = I f_I$
arbitraire	$\longleftarrow \dashv$ R		$e_R = R f_R$
	$\longrightarrow \nearrow$ R		$f_R = e_R / R$

non arbitraire		$e_R = R(f_R)$	
		$f_R = R^{-1}(e_R)$	

Tableau 6.2—Règles d'affectation de la causalité sur les éléments actifs et passifs

Élément	Règle	Loi
	1 seul lien <b>avec</b> trait causal près du 0	$e_1 = e_i, e_2 = e_i,$ $\dots, e_n = e_i$ $a_i f_i = -(a_1 f_1 +$ $a_2 f_2 + \dots + a_n f_n)$
	1 seul lien <b>sans</b> trait causal près du 1	$f_1 = f_i, f_2 = f_i,$ $\dots, f_n = f_i$ $a_i e_i = -(a_1 e_1 +$ $a_2 e_2 + \dots + a_n e_n)$
	affectation symétrique de la causalité	$e_1 = m e_2$ $f_2 = m f_1$ $e_2 = 1/m e_1$ $f_1 = 1/m f_2$
	affectation anti- symétrique de la causalité	$e_1 = r f_2$ $e_2 = r f_1$ $f_1 = 1/r e_2$ $f_2 = 1/r e_1$

Tableau 6.3—Règles d'affectation de la causalité sur les éléments de jonction

On voit que les éléments actifs ont une causalité fixée, les sources d'effort (resp. de flux) imposant de façon externe un effort (resp. un flux). Les éléments résistifs non linéaires ont aussi une causalité fixée. Les processus de stockage d'énergie ont quant à eux une

causalité préférentielle : un élément capacitif produit de préférence un effort et un élément inductif engendre de préférence un flux. Comme les lois de ces éléments relient respectivement la dérivée d'un effort à un flux et la dérivée d'un flux à un effort, on reconnaît là la causalité intégrale adoptée par ailleurs tant par de Kleer et Brown que par Iwasaki et Simon (néanmoins, cette causalité intégrale n'est pas contingente mais seulement préférentielle ; la causalité inversée, dite causalité *dérivée*, où la dérivée d'une variable dépend de cette variable, peut aussi exister). La causalité des éléments résistifs linéaires est par contre en principe dépendante du contexte, c'est-à-dire déterminée par la causalité des autres éléments du système (sources et processus de stockage d'énergie). Si ce n'est pas le cas, elle peut être assignée arbitrairement.

La procédure d'assignation de la causalité (connue sous le nom de SCAP comme « *Sequential Causality Assignment Procedure* ») comprend les étapes suivantes :

1. Affecter la causalité obligatoire aux sources et aux éléments résistifs R non linéaires et répercuter sur l'environnement en respectant les restrictions de causalité aux jonctions ;
2. Affecter une causalité intégrale préférentielle aux éléments I et C et répercuter sur l'environnement en respectant les restrictions de causalité aux jonctions ;
3. Compléter la causalité sur les éléments R linéaires en respectant les restrictions de causalité aux jonctions ;
4. En cas de conflit à une jonction, rechercher l'élément I ou C cause de conflit et lui affecter une causalité dérivée. Boucler à l'étape précédente.

Donnons quelques exemples de cette procédure.

#### **Exemple 6.5**

Appliquons cette procédure à l'exemple du régulateur de pression.

Partons d'une modélisation statique (c'est-à-dire en régime stationnaire, à débit constant) simple, analogue à celle de l'exemple 3.1. Le conduit principal (vanne d'orifice  $S$ ) se modélise comme un composant résistif dont la valeur de la résistance est inversement proportionnelle à  $S$  (cette dépendance est indiquée par une liaison informationnelle), qui donne lieu à une équation de perte de charge :  $Q = cS\sqrt{P_e - P_s}$ , analogue à (éq.3), au sens où l'équation qualitative associée est identique, soit (éq.7). Ce conduit résistif est relié par une jonction 1 (même débit) d'une part à une source d'effort

imposant la pression exogène d'entrée  $P_e$ , dont l'équation est (ég.4), d'autre part à un autre composant résistif pour modéliser la charge en sortie, dont l'équation est (ég.2). En l'absence d'une modélisation dynamique du piston, l'équation de régulation (ég.1) est directement traduite par une liaison de  $P_s$  vers  $S$ . On suppose, en accord avec l'analyse de l'exemple 3.1, que, dans la charge en sortie, le débit impose la pression (on peut montrer que c'est une conséquence du modèle dynamique complet).

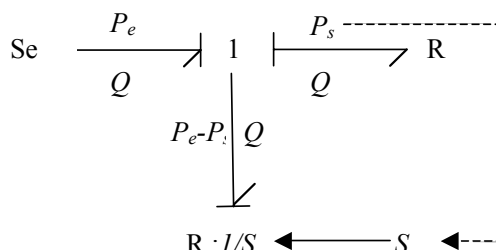


Figure 6.1—Grphe de liaisons du régulateur de pression (modèle statique)

La causalité est alors entièrement déterminée et l'on obtient le graphe de liaisons de la Figure 6.1.

De ce graphe, on déduit la structure causale suivante (Figure 6.2), dont on remarque qu'elle coïncide avec celle établie par la causalité mythique à partir de la propagation d'une perturbation exogène de la pression d'entrée (Figure 3.2), alors même que l'ordonnancement causal d'Iwasaki et Simon n'ordonnait pas causalement les variables  $Q$ ,  $P_s$  et  $S$  (Figure 4.1).

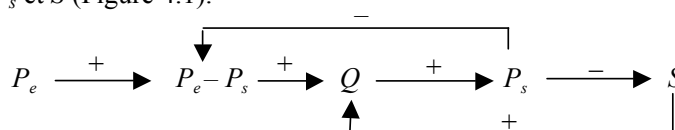


Figure 6.2—Structure causale du régulateur de pression dérivée du graphe de liaisons de la Figure 6.1

### 6.3 Comparaison avec l'ordonnancement causal

Les concepts de causalité issus de la théorie de l'ordonnancement causal et de la méthode des graphes de liaisons s'avèrent être semblables : les ordres causals obtenus sont identiques si les modèles mathématiques sous-jacents sont les mêmes. En effet, du point de vue mathématique, un graphe de liaisons peut être considéré comme un cadre organisationnel qui arrange les équations et les variables de

façon que les substitutions soient représentées par des liens graphiques, à savoir les liaisons énergétiques. La différence fondamentale entre les deux méthodes est que l'ordonnement causal est fondé sur des relations mathématiques abstraites et la méthode des graphes de liaisons sur des principes physiques formels. Le premier peut donc avoir une portée plus générale que les systèmes physiques, mais, lorsqu'on a affaire à ceux-ci, la seconde constitue une représentation plus riche de la connaissance, se prête mieux à l'explication physique et engendre plus d'information causale.

#### Résolution de la causalité

Les graphes de liaisons étant fondés sur un ensemble exhaustif de mécanismes physiques primitifs, ils garantissent que, pour un modèle donné (dont la granularité de détail est fixée), la causalité obtenue aura une résolution maximale, en termes d'explication causale. Cette résolution est automatiquement engendrée, alors qu'elle est laissée au savoir-faire de celui qui écrit les équations du modèle dans le cas de l'ordonnement causal. Par exemple, dans le cas d'une conservation de flux d'énergie, les graphes de liaisons la décomposent en les mécanismes primitifs qui la créent, décomposition qui peut être aisément omise dans le cas de l'ordonnement causal car elle ne découle pas de façon évidente de la structure mathématique du système.

#### Exogénéité

La détermination de l'exogénéité est une question de modélisation (hypothèses sur les interactions possibles entre le système considéré et le reste du monde) et pas de syntaxe. Dans la théorie de l'ordonnement causal, les variables considérées comme exogènes sont les variables du système qui ne sont pas influencées par le système. Les graphes de liaisons distinguent exogénéités directe et indirecte mais cette distinction ne tient qu'à la distinction entre variables et paramètres (constants) et peut tout aussi bien être faite avec l'ordonnement causal.

#### Complétude

La théorie de l'ordonnement causal requiert que le modèle mathématique (statique, dynamique ou mixte) soit complet. Mais la vérification de la complétude et la réparation du modèle en cas d'incomplétude revient à la personne qui modélise le système. Par contre, si le modèle est exprimé sous forme d'un graphe de liaisons, l'assignation de la causalité vérifie automatiquement la complétude. En effet, elle assure que chaque variable d'effort ou de flux est le résultat d'une unique expression dans laquelle cette variable n'intervient pas. En phase de modélisation, il est toujours possible de construire un système qui ne soit pas complet. Mais la procédure SCAP détectera immédiatement le problème et le localisera en termes d'erreurs causales.

### Rétroaction

Dans la théorie de l'ordonnement causal, une rétroaction est détectée quand l'analyse aboutit à un système d'équations qui ne peuvent être résolues que simultanément. Aucune information n'est fournie sur le type ou le signe de la rétroaction. Avec les graphes de liaisons, le type et le signe découlent automatiquement du graphe causalement augmenté. De plus la présence d'une rétroaction algébrique est détectée dès l'application de SCAP qui ne préconise pas dans ce cas d'orientation causale préférée dans la boucle. De façon similaire à la causalité mythique qui produit toutes les directions causales, une direction causale arbitraire est assignée à l'intérieur de la boucle de rétroaction, au contraire de l'ordonnement causal.

### 7. CONCLUSION

Le domaine du Raisonnement Qualitatif s'attache à prédire et expliquer le comportement des mécanismes physiques en termes qualitatifs. Ceci sous-entend de pouvoir produire une prédiction qualitative du comportement d'un système physique mais aussi d'être capable de construire une explication causale du pourquoi de ce comportement qualitatif. C'est pourquoi la causalité et les modèles causals tiennent une place importante dans ce domaine. La causalité est vue comme une composante du raisonnement de sens commun et les modèles causals permettent d'exploiter une causalité opérationnelle à des fins d'explication ou de prédiction.

Cet article a présenté comment la causalité est abordée d'un point de vue conceptuel ainsi que les différentes approches opérationnelles permettant de produire la *structure causale* (ou *graphe d'influences*) d'un système à partir d'une connaissance issue des principes premiers régissant le comportement des composants (ou mécanismes) du système, i.e. un modèle équationnel. Les trois principales approches (causalité mythique, ordonnancement causal, procédure SCAP des graphes de liaisons) ont été présentées<sup>3</sup> et produisent des structures causales identiques qui reflètent bien la façon dont l'ingénieur appréhende les relations de cause à effet. Elles se basent pourtant sur des concepts très éloignés, ce qui les conduit à

---

<sup>3</sup> A côté de ces trois approches, d'autres approches moins connues ont été proposées. Mentionnons l'ordonnement causal fondé sur une analyse dimensionnelle des quantités physiques (Shen et al., 1999) et le Graphe Informationnel Causal (GIC) qui est proposé par Hautier et Faucher (1996) comme un modèle de connaissance mettant en œuvre les lois de la physique dans le strict respect du principe de causalité. Il est à rapprocher du concept de Graphe de Liaisons mais l'idée de causalité dérivée y est exclue puisqu'allant, selon les auteurs, à l'encontre des systèmes physiquement réalisables.



une interprétation des rétroactions très différente qui a fait l'objet d'une forte controverse rapportée en section 5.

Les arguments avancés par les auteurs nous amènent à conclure que les différences apparentes reflètent en réalité des différences dans les points de vue de modélisation. Ceci met en relief l'importance de la modélisation. En effet, les structures causales produites souffrent d'*instabilité contextuelle*, comme le remarque Skorstad (1990 ; 1992), c'est-à-dire qu'elles dépendent de l'ensemble de variables considérées comme exogènes et du niveau d'abstraction temporelle adopté. C'est l'environnement dans lequel le modélisateur place le système qui définit les variables exogènes. Il ne faut donc pas négliger la connaissance de cet environnement, à mettre en vis-à-vis avec la connaissance du système lui-même et notamment des différentes échelles de temps auxquelles se déroulent les différents processus. Ces processus définissent quant à eux l'ensemble des variables quasi-exogènes qui jouent un rôle tout aussi déterminant dans la causalité que les variables exogènes.

Le niveau d'approximation du modèle joue de même un rôle primordial. Il faut noter que l'écriture d'un graphe de liaisons peut s'avérer délicate pour un modélisateur peu pénétré des lois physiques du domaine considéré. Ceci pourrait paraître surpassé par l'approche de l'ordonnement causal qui est totalement automatique. Il ne faut cependant pas s'y tromper : cette approche demande à ce que les équations à partir desquelles la méthode s'applique soient structurelles, c'est-à-dire que chacune d'elles doit effectivement représenter le comportement d'un mécanisme de base. Le principe de localité est donc dans les équations, et l'obtention d'un tel modèle demande sans aucun doute des connaissances certaines sur les lois physiques régissant les phénomènes.

### Références

- Bandekar, V. (1990). Causal structures: Computations and Applications, in *Proc. of the 9th European Conference on Artificial Intelligence ECAI-90*, Stockholm, Suède, pp. 71-76.
- Brown, A. (1976). Qualitative knowledge, causal reasoning and the localization of failures, *Technical report 362, MIT Artificial Intelligence Laboratory*, Cambridge, MA, USA.
- Dauphin-Tanguy, G. (2000). *Les bond graphs*, traité IC2, Hermès.
- Davis, R. (1983). Diagnosis via causal reasoning: paths of interaction and the locality principle, in *Proc. of the 3rd National Conference on Artificial Intelligence AAAI-83*, Washington, DC, USA, pp. 88-94 ; aussi in D. Weld, et J. de Kleer, (1990), *Readings in Qualitative Reasoning About Physical Systems*, Morgan Kaufmann, pp. 535-541.

- de Kleer, J. (1977). Multiple Representations of Knowledge in a Mechanics Problem-Solver, in *Proc. of the 5th International Joint Conference on Artificial Intelligence IJCAI-77*, Cambridge, MA, pp. 299-304.
- de Kleer, J. (1979). The origin and resolution of ambiguities in causal arguments, in *Proc. of the 6th International Joint Conference on Artificial Intelligence IJCAI-79*, Tokyo, Japon, pp. 197-204 ; aussi in D. Weld, et J. de Kleer, (1990), *Readings in Qualitative Reasoning About Physical Systems*, Morgan Kaufmann, pp. 624-630.
- de Kleer, J., Brown, J.S. (1984). A qualitative physics based on confluences, *Artificial Intelligence* 24, pp. 7-83 ; aussi in J. Hobbs, et R. Moore, (1985), *Formal Theories of the Commonsense World*, Ablex, Norwood, pp. 109-183 ; aussi in D. Weld, et J. de Kleer, (1990), *Readings in Qualitative Reasoning About Physical Systems*, Morgan Kaufmann, pp. 88-125.
- de Kleer, J., Brown, J.S. (1986). Theories of causal ordering, *Artificial Intelligence* 29, pp. 33-62.
- Forbus, K.D. (1984). Qualitative process theory, *Artificial Intelligence* 24, pp. 85-168 ; aussi in D. Weld, et J. de Kleer, (1990), *Readings in Qualitative Reasoning About Physical Systems*, Morgan Kaufmann, pp. 178-219.
- Forbus, K.D., Gentner, D. (1983). Causal reasoning about quantities, in *Proc. of the fifth Annual Conference of the Cognitive Science Society*, Lawrence Erlbaum and Associates, pp. 196-206 ; aussi in D. Weld, et J. de Kleer, (1990) *Readings in Qualitative Reasoning About Physical Systems*, Morgan Kaufmann, pp. 666-677.
- Gondran, M., Minoux, M. (1979). Graphes et algorithmes, Eyrolles, Paris.
- Hautier, J.P., Faucher, J. (1996). Le Graphe Informationnel Causal, *Bulletin de l'Union des Physiciens* 90, juin 1996, pp. 167-189.
- Iwasaki, Y. (1988). Causal ordering in a mixed structure, in *Proc. of the 7th National Conference on Artificial Intelligence AAAI-88*, Saint Paul, Minnesota, USA, pp. 313-318.
- Iwasaki, Y., Simon, H.A. (1986a). Causality in device behavior, *Artificial Intelligence* 29, pp. 3-32.
- Iwasaki, Y., Simon, H. (1986b). Theories of causal ordering: reply to de Kleer and Brown, *Artificial Intelligence* 29, pp. 63-68 ; aussi in D. Weld, et J. de Kleer, (1990), *Readings in Qualitative Reasoning About Physical Systems*, Morgan Kaufmann, pp. 661-665.
- Iwasaki, Y., Simon, H. (1993). Retrospective on "Causality in device behavior", *Artificial Intelligence* 59, pp. 141-146.
- Iwasaki, Y., Simon, H.A. (1994). Causality and model abstraction, *Artificial Intelligence* 67, pp. 143-194.
- Karnopp, D., Rosenberg, R. (1975). *System dynamics: a unified approach*, John Wiley & Sons.
- Lee, M., Compton, P. (1994). Context-dependent causal explanations, in *Proc. of the 8th International Workshop on Qualitative Reasoning about Physical Systems QR'94*, Nara, Japon, pp. 176-186.

- Mosterman, P. J. (1997). *Hybrid dynamic systems: A hybrid bond graph modeling paradigm and its application in diagnosis*, PhD dissertation, Vanderbilt university, Nashville, TN, USA.
- Mosterman, P. J., Biswas, G. (1997). Formal specifications for hybrid dynamical systems, in *Proc. of the 15th International Joint Conference on Artificial Intelligence IJCAI-97*, Nagoya, Japon, pp. 568-573.
- Mosterman, P.J., Manders, E.-J., Biswas, G. (2000). Qualitative dynamic behavior of physical system models with algebraic loops, in *Proc. of the 11th International Workshop on Principles of Diagnosis DX'00*, Morelia, Mexique.
- Paynter, H. (1961). *Analysis and design of engineering systems*, MIT Press.
- Porcheron, M., Ricard, B. (2000). An application of abductive diagnostic methods to a real world problem, in *Proc. of the 11th International Workshop on Principles of Diagnosis DX'00*, Mont St Michel, France, pp. 87-94.
- Porté, N., Boucheron, S., Sallantin, S., Arlabosse, F. (1988). An algorithmic view at causal ordering, in *Proc. of 2nd International Workshop on Qualitative Physics QR'88*, Paris, France.
- Rose, P., Kramer, M.A. (1991). Qualitative analysis of causal feedback, in *Proc. of the 9th National Conference on Artificial Intelligence AAAI-91*, Anaheim, CA, USA, pp. 817-823.
- Rosenberg, R., Karnopp, D. (1983). *Introduction to physical system dynamics*, series in mechanical engineering, Mac Graw Hill.
- Shen, Q., Peng, T., Milne, R. (1999). Dimensional analysis based causal ordering, in *Proc. of the 13<sup>th</sup> International Workshop on Qualitative Reasoning QR'99*, Loch Awe, Ecosse, pp. 193-202.
- Simon, H.A. (1952). On the definition of the causal relation, *The Journal of Philosophy*, 49, pp. 517-528 ; aussi in H.A. Simon (1977), *Models of Discovery*, Reidel, pp. 86-91.
- Simon, H.A. (1953). Causal ordering and identifiability, in W.C. Hood, et T.C. Koopmans, *Studies in Econometric Models*, Wiley, pp. 49-74 ; aussi in H.A.Simon (1977), *Models of Discovery*, Reidel, pp. 53-70.
- Skorstad, G. (1990). Clustered causal ordering, in *Proc. of the 4th International Workshop on Qualitative Physics QR'90*, Lugano, Suisse, pp. 290-299.
- Skorstad, G. (1992). Finding stable causal interpretations of equations, in P. Struss (Ed.), *Recent Advances in Qualitative Physics*, pp. 399-412.
- Tomlin, C.J., Greenstreet, M.R. (Eds.) (2002). *Hybrid Systems: Computation and Control*, in *Proc. of the 5<sup>th</sup> International Workshop HSCC 2002*, Stanford, CA, USA.
- Top, J., Akkermans, H. (1991). Computational and physical causality, in *Proc. of the 12th International Joint Conference on Artificial Intelligence IJCAI-91*, Sydney, Australie, pp. 1171-1176.
- Travé-Massuyès, L., Dague, P. (2003). *Modèles et raisonnements qualitatifs*, chapitre 7 : *Raisonnement causal en physique qualitative*, pp. 207-268, traité IC2, Hermès Science, Lavoisier, ISBN 2-7462-0744-3.

- Travé-Massuyès, L., Pons, R. (1997). Causal ordering for multiple mode systems, in *Proc. of the 11<sup>th</sup> International Workshop on Qualitative Reasoning QR'97*, Cortona, Italie, pp. 203-214.
- Weld, D., de Kleer, J. (Eds.) (1990). *Readings in Qualitative Reasoning About Physical Systems*, Morgan Kaufmann, San Mateo, CA, USA.
- Williams, B.C. (1984). Qualitative analysis of MOS circuits, *Artificial Intelligence* 24, pp. 281-346.
- Williams, B.C. (1990). Temporal qualitative analysis: explaining how physical systems work, in D. Weld, et J. de Kleer (1990), *Readings in Qualitative Reasoning About Physical Systems*, Morgan Kaufmann, pp. 133-177.