

Integrating Document-Based and Knowledge-Based Models for Clinical Guidelines Analysis



Gersende Georg^{1,2,3} and Marc Cavazza⁴

¹INSERM U 872, Eq. 20, SPIM

²Université Paris Descartes, UMR S 872

³Centre de Recherche des Cordeliers, Université Pierre et Marie Curie - Paris 6,
UMR S 872, Paris

⁴School of Computing, University of Teesside

Outline

- Context of the work
- Technical approach
 - Integration of knowledge-based and document-based approaches
- Hierarchical Task Planning (HTN)
- G-DEE
 - Environment dedicated to the study of Clinical Guidelines incorporating text processing functions
- Conclusions

Computerization of Clinical Guidelines

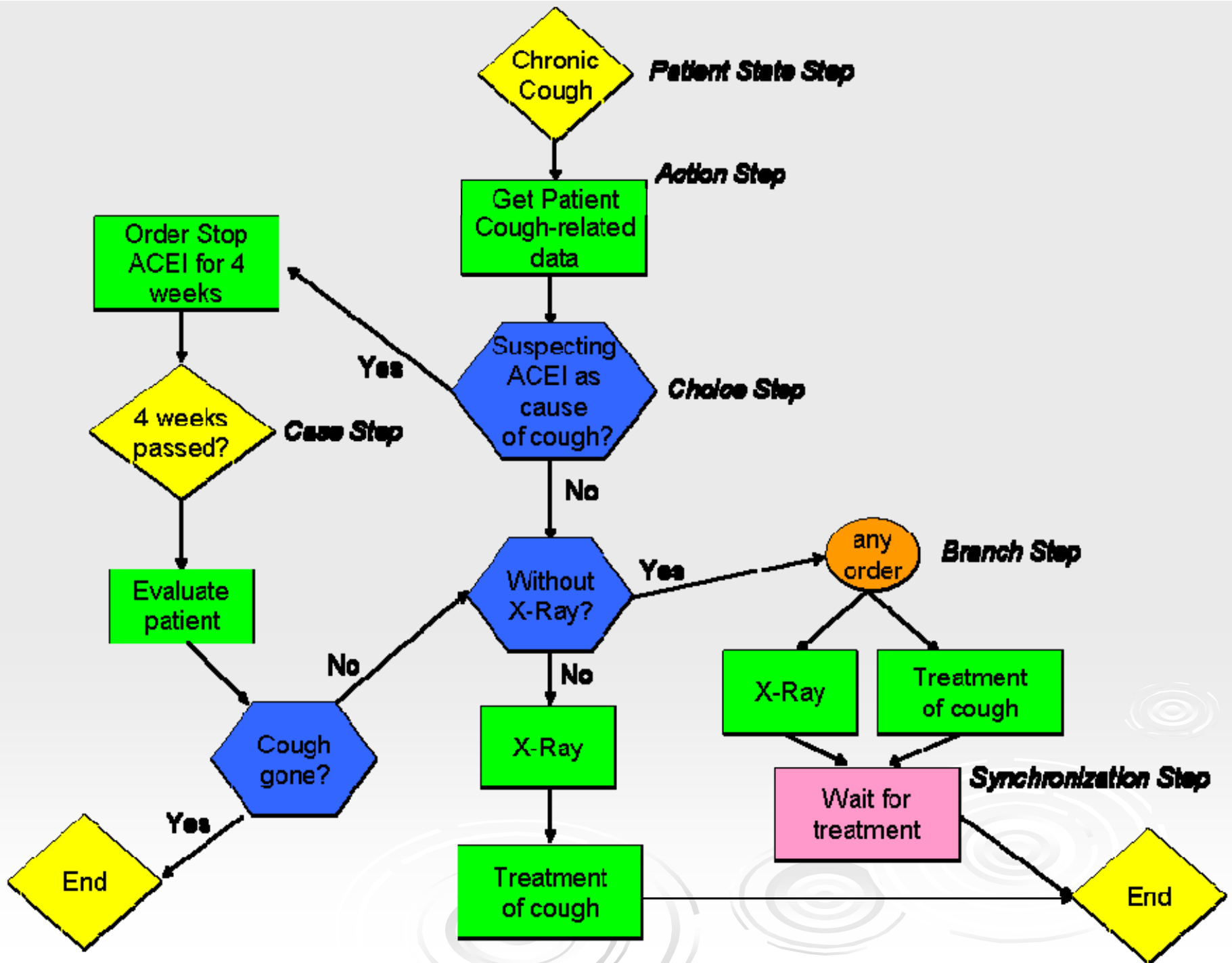
Translation of guideline into computer-executable statements is complex [Shiffman et al, 1999]

Approaches centered on knowledge formalization

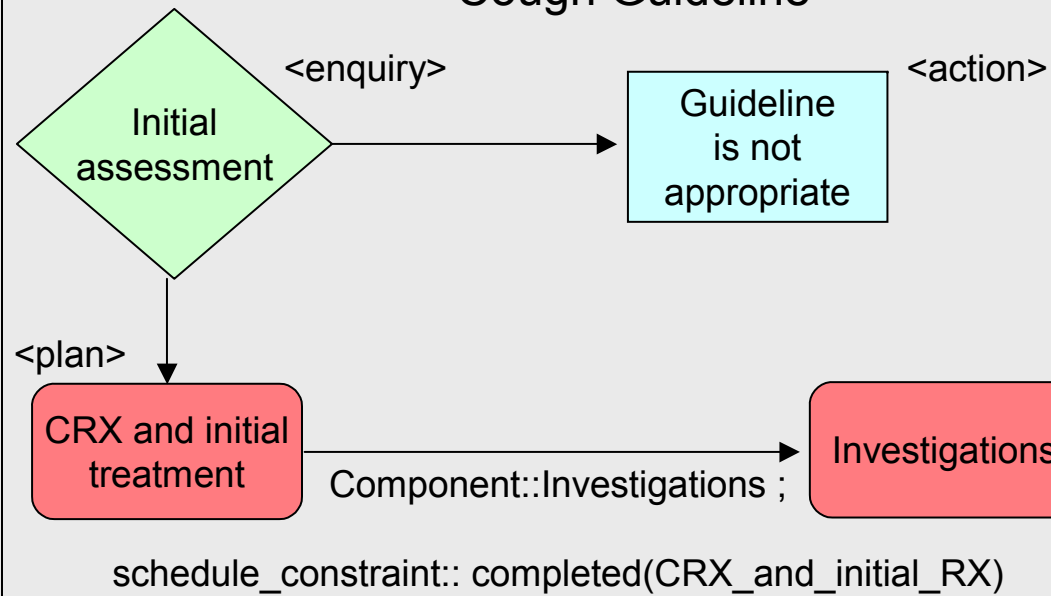
- Description of decision process that contents in clinical guidelines
 - GLIF [Greenes et al, 1999]
 - PRO*forma* [Fox & Rahmanzadeh, 1998]
 - GUIDE [Ciccarese et al, 2003]
 - PRODIGY [Purves et al, 1999]
- Aimed at elaborating decision systems

Approaches centered on document engineering

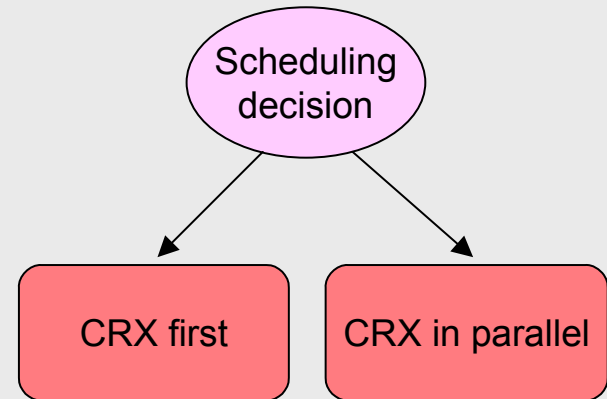
- Methods using document as a representation support
 - HGML [Hagerty et al, 2000]
 - GEM [Shiffman et al, 2000]
- Aimed at managing the document



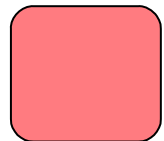
Cough Guideline



CRX and initial treatment



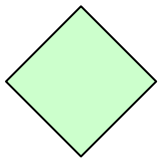
Keys



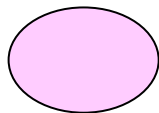
Plans



Actions



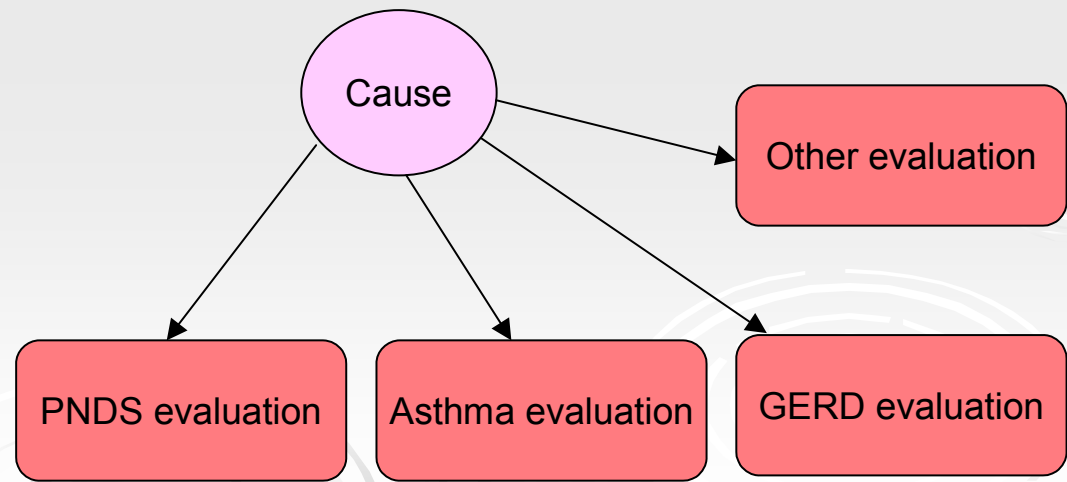
Enquiries

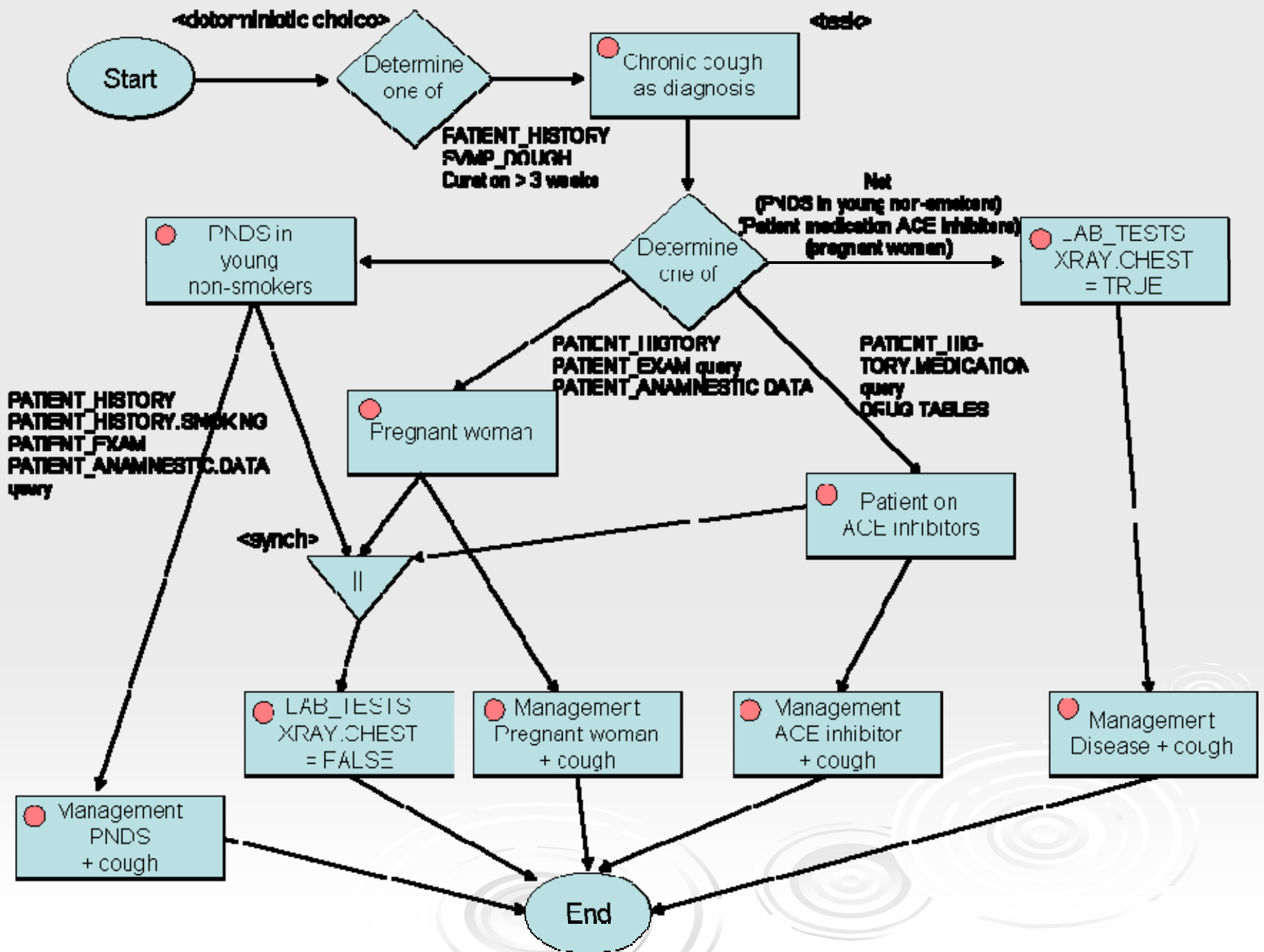


Decisions

→ Scheduling constraints

Investigations





Document-based Approaches

- Adding information on text using marking-up
- GEM is the most known model and has been recognized as a standard to structure clinical guidelines
- Translation of text to document models faces limitations
 - Documents are subject to variations in style
 - Interpretation problems affecting encoding

Comparison of Clinical Guidelines Encoding

Knowledge-based approach

- Formalisms represent different processes : decompositions, temporal and conditional aspects
- Divided into algorithmic representations, plans and Petri nets
- Document must be analyzed and understood as a whole to be translated into formalisms
- Can directly be integrated into Decision Support Systems (DSS)

Document-based approach

- Enhancement of text by elements of information
- Requires an additional level of interpretation to extract rules or decision trees
- A pivot representation for ulterior transformations or the generation of others representations

-
- Extraction of complete knowledge structures from free text using Information Extraction techniques
 - HGML (Hagerty et al. 2006), LASSIE (Kaiser et al. 2005)

Approach

- Integrate knowledge-based and document-based approaches.
- Test case: study the structure of clinical guidelines (with the objective of assessing their consistency and completeness)
- Analyze textual guidelines using a software environment based on
 - guideline document encoding using mark-ups
 - protocol representation using the most general formalism

Part I

A General Formalism Approach

Modeling Guidelines with Planning Formalisms

- Knowledge-based approaches are based on knowledge structures centered on clinical actions
 - PROforma (Fox et al. 1996)
 - Prodigy (Johnson et al. 2000)
- related to action representations encountered as part of planning formalisms such as STRIPS or PDDL
- Role of planning formalisms and their applicability to clinical guidelines modeling (Bradbook et al. 2005)

Using HTN to represent Guidelines Protocols

- Hierarchical Task Networks (HTN) are one of the most successful planning formalisms and focus on tasks and knowledge
- HTN are used in a variety of implemented systems
 - Robotics, game playing (bridge), virtual characters animation
- HTN are best applied to knowledge-intensive Planning problems
- Their top-down descriptions are well suited to the description of clinical protocols

HTN Formalization of Clinical Guidelines

- Appropriate to represent multi-step decomposable processes, provided they can be decomposed into independent sub-tasks
- The explicit nature of AND/OR graphs allows a direct visualization that also facilitates knowledge elicitation
- Various steps of clinical care can be represented as an AND/OR graph, in essence formalizing an *explicit* HTN
 - The main task (the overall goal of the clinical protocol) has been decomposed *a priori* and entirely, down to the level of grounded actions
 - Each sub-task should be associated pre-conditions as well as post-conditions
 - Heuristics guide node selection at the level of OR nodes, and costs associated to the actual clinical actions

Part II

A Document Engineering Environment

G-DEE (Guidelines Document Engineering Environment)

- A software environment for the study of Clinical Guidelines that incorporates automatic text processing functions
- Automatically performs XML encoding of guidelines based on the recognition of the guideline's linguistic content
- These processing functions recognize specific natural language expressions corresponding to the linguistic formulation of elementary recommendations (*deontic operators*)

Analyse de la sélection

Analyse du texte

Phrase à modifier :

En cas de **signe évocateur ou d'antécédent d'infection urinaire**, **il est recommandé de pratiquer** un ECBU (accord professionnel) (1).

A

Modification Front Scope

Modification Back Scope

Validation

☒ Front Scope☒ Back Scope

Règle

SI **signe évocateur ou d'antécédent d'infection urinaire**ALORS **pratiquer**

un ECBU (accord professionnel)

D

Texte à analyser

III.2. Examens complémentaires.

Il est proposé de vérifier la stérilité des urines par bandelette urinaire. En cas de **signe évocateur ou d'antécédent d'infection urinaire**, **il est recommandé de pratiquer** un ECBU (accord professionnel) (1).

Le dosage de la créatininémie n'est pas proposé à titre systématique. Il n'est recommandé que chez les patients présentant des facteurs de risque d'insuffisance rénale, celle-ci n'étant affirmée que sur le calcul de la clearance de la créatinine (accord professionnel).

La débitmétrie urinaire n'est pas proposée en routine dans le bilan initial d'une HBP symptomatique. C'est un examen optionnel pratiqué en milieu spécialisé (accord professionnel).

L'échographie de l'arbre urinaire par voie abdominale n'est pas proposée à titre systématique dans le bilan initial de l'HBP symptomatique (accord professionnel). Cet examen peut être utile pour le diagnostic de vessie de lutte, de calcul vésical ou de dilatation du haut appareil. Il a été démontré que la mesure du résidu postmictionnel et du volume prostatique par échographie sus-pubienne n'était pas fiable.

Le bilan urodynamique n'est pas recommandé systématiquement dans le bilan initial de l'HBP symptomatique. Cet examen invasif peut être utile en cas de comorbidité, notamment neurologique et pour préciser les indications thérapeutiques en milieu spécialisé (accord professionnel).

Une échographie prostatique par voie transrectale n'est pas recommandée lors du bilan initial de l'HBP symptomatique. Cet examen n'a aucune place dans le diagnostic, le bilan ni la surveillance d'une HBP

Document GEM

<Decision.variable> **En cas de signe évocateur ou d'antécédent d'infection urinaire**
</Decision.variable>

<OpReco> **il est recommandé de pratiquer** </OpReco>

<Action> un ECBU (accord professionnel) </Action>

C

Linguistic Analysis

- *Deontic propositions* are the most characteristic linguistic structures of normative texts [Kalinowski¹]
 - Previous work on legal texts by Moulin et al.²
- These are based on deontic operators such as (in French):
 - *pouvoir* (to be allowed to or may),
 - *devoir* (should or ought to),
 - *interdire* (to forbid)
- Frequency of deontic verbs for a set of 20 French Clinical Guidelines (composed of 83,997 word occurrences)³

¹ Kalinowski G. La Logique Dédutive. Presses Universitaires de France (in French); 1996.

² Moulin B, Rousseau D. Knowledge acquisition from prescriptive texts. ACM, 1990: 1112:1121.

³ Georg G, Colombet I, Jaulent MC. Structuring Clinical Guidelines through the Recognition of Deontic Operators. Stud Health Technol Inform. 2005;116:151-6.

Grammar of Deontic Operators

- Identification of specific syntactic structures for deontic operators
- Extended corpus of 17 documents to maximize syntactic coverage (clinical guidelines / consensus conferences / medical teaching material)
- Deontic operators (for French):
 - *recommander* (to recommend)
 - *conseiller* (to advise)
 - *préférer* (to prefer)
 - *envisager* (to consider)
 - *éviter* (to avoid)
 - *prescrire* (to prescribe)
 - *proposer* (to propose)
 - *traiter* (to treat)

Syntactic formalism

- Finite-State Transition Networks¹ (FSTN)
- Convenient way to define specialized structures, including their morphological variants
- 170 syntactic patterns described from our corpus analysis (corresponding to 65 deontic operators)*
- 12,000 FSTN* once morphological variants are taken into account

¹ Roche E, Schabes Y. *Finite-State Language Processing*. MIT Press; 1997.

* Latest version

Scopes of Deontic Operators

- Scopes correspond to the operands of deontic operators
 - Front-Scope: a scope that precedes a deontic operator
 - Back-Scope: a scope that follows the operator

<FrontScope> Investigation for secondary hypertension (with specific laboratory tests or imaging)
</FrontScope> should be considered <BackScope>
in young hypertensive patients (under 30 years old)
</BackScope>.

- The encoding can serve as a basis for further processing
 - The extraction of decision rules under textual format
 - The encoding using GEM categories

VI. TRAITEMENT PHARMACOLOGIQUE (150).

VI.1. Choix des antihypertenseurs (151).

La réduction du risque cardio-vasculaire est avant tout dépendante de la baisse de la pression artérielle, quelle que soit la classe d'antihypertenseur utilisée (grade A) (86).

Dans l'HTA essentielle non compliquée, les 5 classes d'antihypertenseurs majeurs [les diurétiques thiazidiques, les bêta-bloquants, les inhibiteurs calciques, les inhibiteurs de l'enzyme de conversion (IEC) et les antagonistes des récepteurs de l'angiotensine II (ARA-II)] ont montré un bénéfice sur la morbi-mortalité cardio-vasculaire, dans les essais cliniques (87).

Ces 5 classes d'antihypertenseurs **peuvent donc être proposées** en première intention dans la prise en charge d'un hypertendu essentiel non compliqué (grade A) (88). Le choix d'un traitement médicamenteux **sera adapté** à chaque patient en fonction des **indications préférentielles de certaines classes dans des situations cliniques articulières** (cf (89).

infra) (en accord avec les études cliniques), de l'efficacité et de la tolérance des médicaments déjà pris par le patient, de l'existence de comorbidités pouvant justifier ou contre-indiquer certains antihypertenseurs, et du coût du traitement et de sa surveillance, en sachant que le diurétique thiazidique fait partie des classes dont le coût journalier est le plus faible (90).

En cas de **prescription d'un traitement par IEC ou ARA-II**, **il est recommandé de prescrire** un bilan associant kaliémie et créatininémie dans un délai de 7 à 15 jours, après le début du traitement (91).

Si la **créatininémie s'élève de plus de 20 à 30 %**, **il est recommandé d'arrêter** l'IEC ou l'ARA-II et de demander un avis spécialisé (92). Pour favoriser l'observance, une prise unique quotidienne (monoprise) utilisant un médicament de longue durée d'action **sera préférée** (93).

A partir d'études menées chez des patients afro-américains et africains des Caraïbes, il apparaît que l'HTA chez ces patients, est plus sensible aux diurétiques et aux inhibiteurs calciques qu'aux classes thérapeutiques agissant sur le système rénine-angiotensine (IEC, ARA-II) ou qu'aux bêta-bloquants (94).

VI.2. Stratégie d'adaptation du traitement médicamenteux (152).

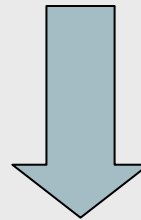
Il est recommandé de débiter par une monothérapie (95). Une association fixe d'antihypertenseurs à doses faibles, ayant l'AMM en première intention pour l'indication HTA, **peut également être proposée** (96).

En deuxième intention, une bithérapie **sera instaurée** dans un délai d'au moins 4 semaines, en cas de **réponse tensionnelle insuffisante au traitement initial** (97).

Une bithérapie **pourra être instaurée** dans un délai plus court, dans les cas suivants chez le patient ayant une PA = 180-110 mmHg quel que soit le nombre de facteurs de risque cardio-vasculaire associés, chez le patient ayant une PA de 140-179/90-109 mmHg et à risque cardio-vasculaire élevé (98).

Automatic Structuring in the GEM Format

A `<FrontScope>` Investigation for secondary hypertension (with specific laboratory tests or imaging) `</FrontScope>` **should be considered** `<BackScope>` in young hypertensive patients (under 30 years old) `</BackScope>`.



B

Document GEM

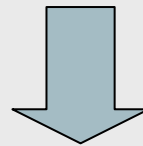
`<Decision.variable>` in young hypertensive patients (under 30 years old) `</Decision.variable>`

`<OpReco>` **should be considered** `</OpReco>`

`<Action>` Investigation for secondary hypertension (with specific laboratory tests or imaging) `</Action>`

Automatic Structuration in the Decision Rule Format

Ⓐ <FrontScope> Investigation for secondary hypertension (with specific laboratory tests or imaging) </FrontScope> **should be considered** <BackScope> in young hypertensive patients (under 30 years old) </BackScope>.



Règle

IF young hypertensive patients (under 30 years old)

THEN **consider**

Investigation for secondary hypertension (with specific laboratory tests or imaging)

Part III

G-DEE & HTN

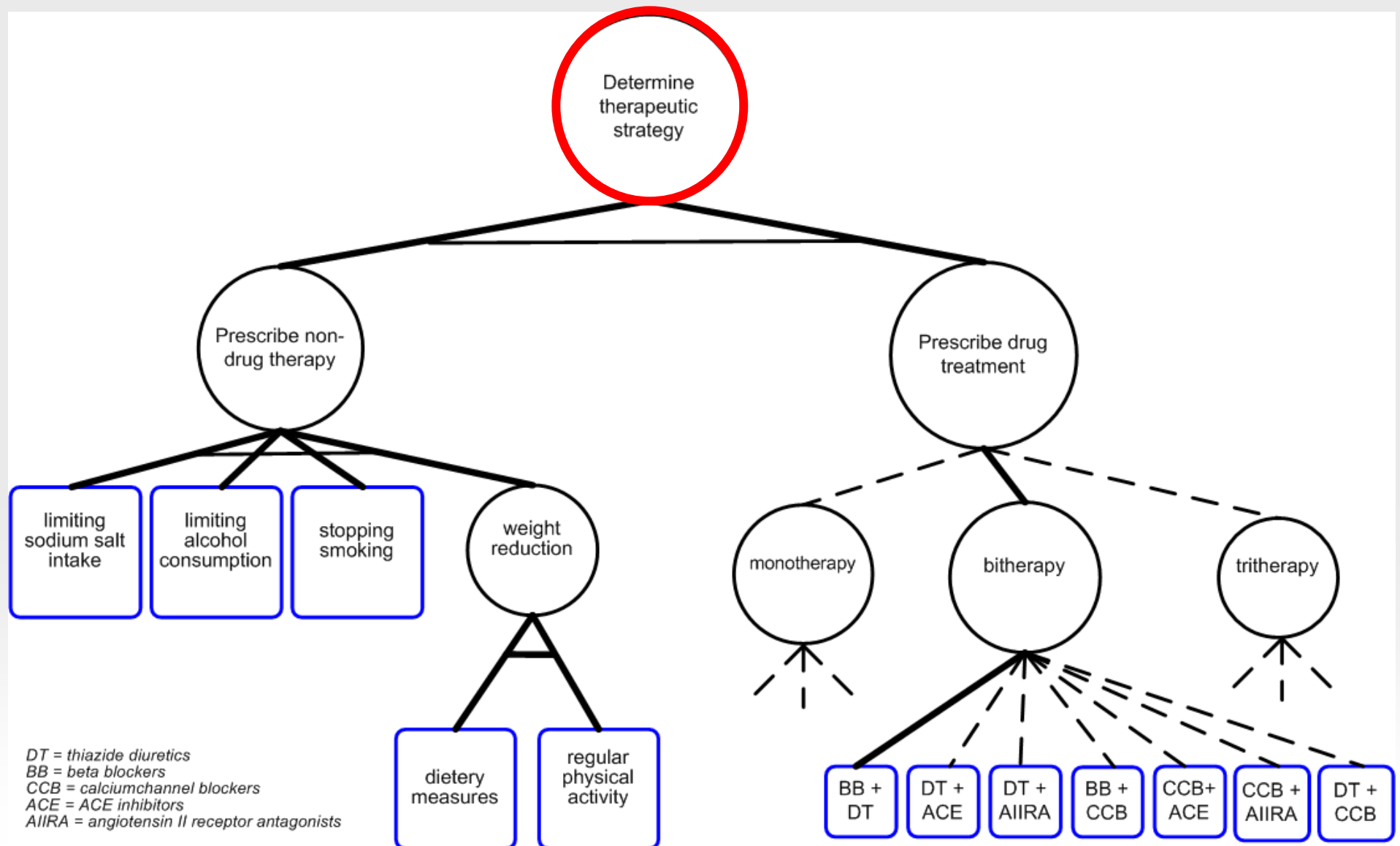
System Overview

- An experimental software platform integrating
 - A document engineering environment
 - An HTN-based module
- The protocol has been modeled as an HTN and the system-driven exploration of its AND/OR graph drives the interactive consultation of the guideline document model to assist in the instantiation of the solution sub-graph
- Synchronization is based on the automatic extraction of information from the marked-up guideline corresponding to the HTN sub-task

Key Ideas

- Instantiate HTN using the contents of guidelines, more specifically recommendations that constitute essential knowledge
- Aims at obtaining a sub-graph representing an instance of the guidelines recommendations on specific patient data
- Our algorithm for traversing HTN and guideline is based on
 - “Offline heuristic calculation mode”
 - “Online heuristic mode”

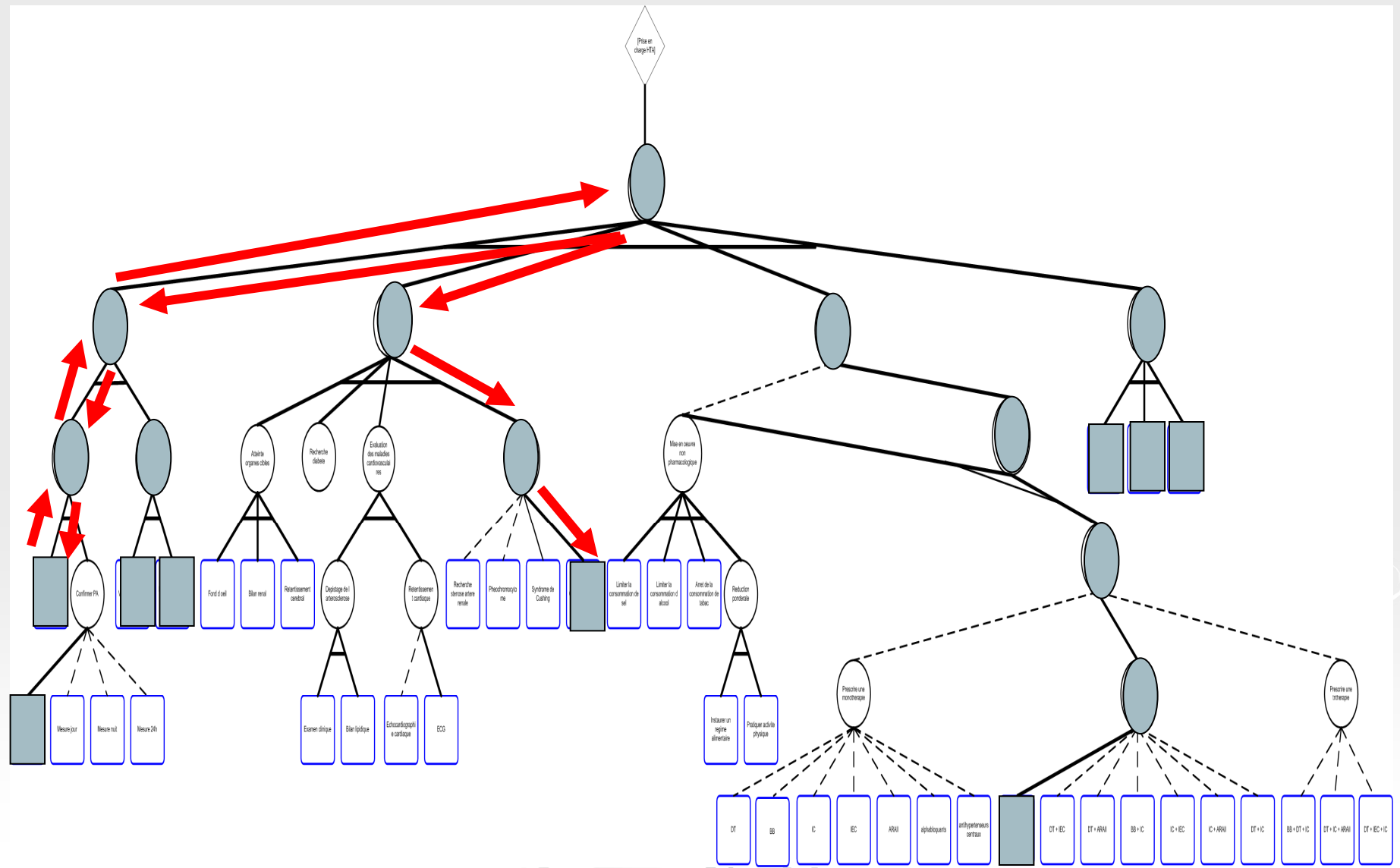
Example: the Hypertension HTN



Overview of the Process

- The HTN is represented as an AND/OR graph traversed using AO*
- The user interprets the different recommendations highlighted by G-DEE in the dedicated interactive window, and is able to specify the corresponding heuristic value for the node considered
- The final output of an interactive session is an explicit task decomposition (only containing AND nodes) which is ready for instantiation on the data specific to a patient profile

Description of AO*



Example Run:

French Hypertension Guidelines

- The overall system behavior consists in traversing the HTN from the top node and extracting a solution graph
- Data are accessed for an example patient, which will drive the instantiation of the various pre-condition of the plan operators
- Determine which textual data is instantiating operators and which one is used by heuristic rules
 - e.g., *age* in the exploration of secondary hypertension

Analyse du texte

[illegible]

Modification Black Scripts

 [Facebook](#)

Pour favoriser l'observance, une prise unique quotidienne (monoprise) utilisant un médicament de longue durée

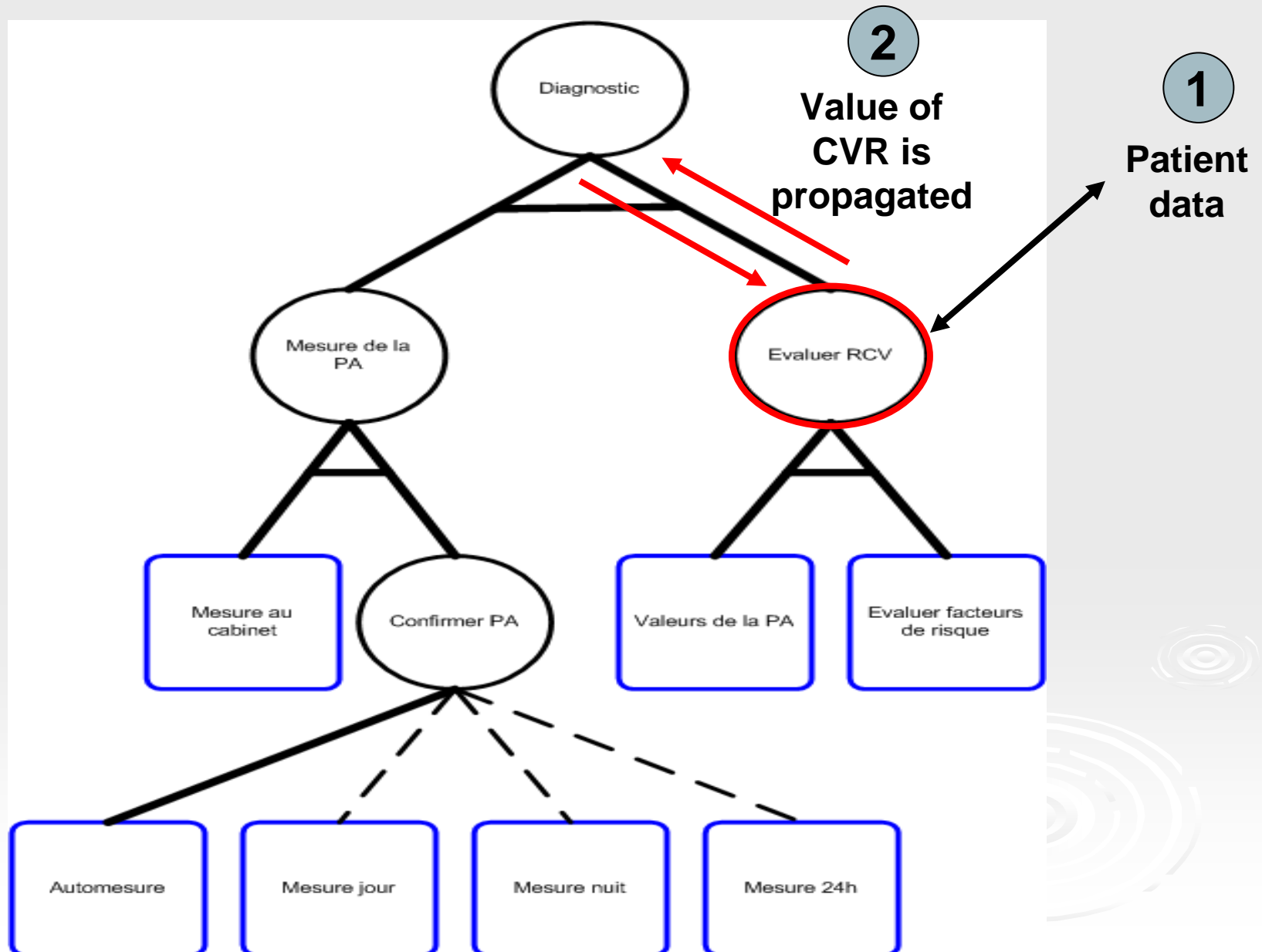
A

Synchronization of HTN Traversal and Document Exploration

- An HTN for the management of hypertension¹ has been developed independently of the textual guideline itself (by using knowledge from other sources, e.g. textbooks, experts)
- Synchronization should relate the contents of HTN nodes traversed (concepts) to the contents of the textual guidelines

¹“Management of adults with essential hypertension – 2005 update” (<http://www.has-sante.fr>)

Example of Determining Heuristic in the Offline Mode



Texte à analyser

III.2. Examens complémentaires.

Il est proposé de vérifier la stérilité des urines par bandelette urinaire. **En cas de signe évocateur ou d'antécédent d'infection urinaire, il est recommandé de pratiquer un ECBU (accord professionnel)** (1).

Le dosage de la créatininémie n'est pas proposé à titre systématique. Il n'est recommandé que chez les patients présentant des facteurs de risque d'insuffisance rénale, celle-ci n'étant affirmée que sur le calcul de la clearance de la créatinine (accord professionnel).

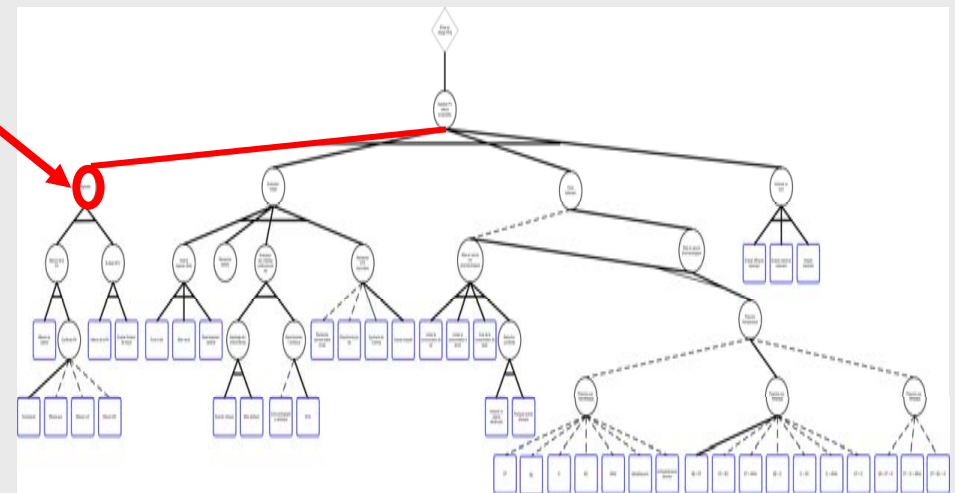
La débitmétrie urinaire n'est pas proposée en première intention dans le bilan initial d'une HBP symptomatique. C'est un examen optionnel pratiqué en milieu spécialisé (accord professionnel).

L'échographie de l'arbre urinaire par voie abdominale n'est pas proposée à titre systématique dans le bilan initial de l'HBP symptomatique (accord professionnel). Cet examen peut être utile pour le diagnostic de vessie de lutte, de calcul vésical ou de dilatation du haut appareil. Il a été démontré que la mesure du résidu postmictionnel et du volume prostatique par échographie sus-pubienne n'était pas fiable.

Le bilan urodynamique n'est pas recommandé systématiquement dans le bilan initial de l'HBP symptomatique. Cet examen invasif peut être utile en cas de comorbidité, notamment neurologique et pour préciser les indications thérapeutiques en milieu spécialisé (accord professionnel).

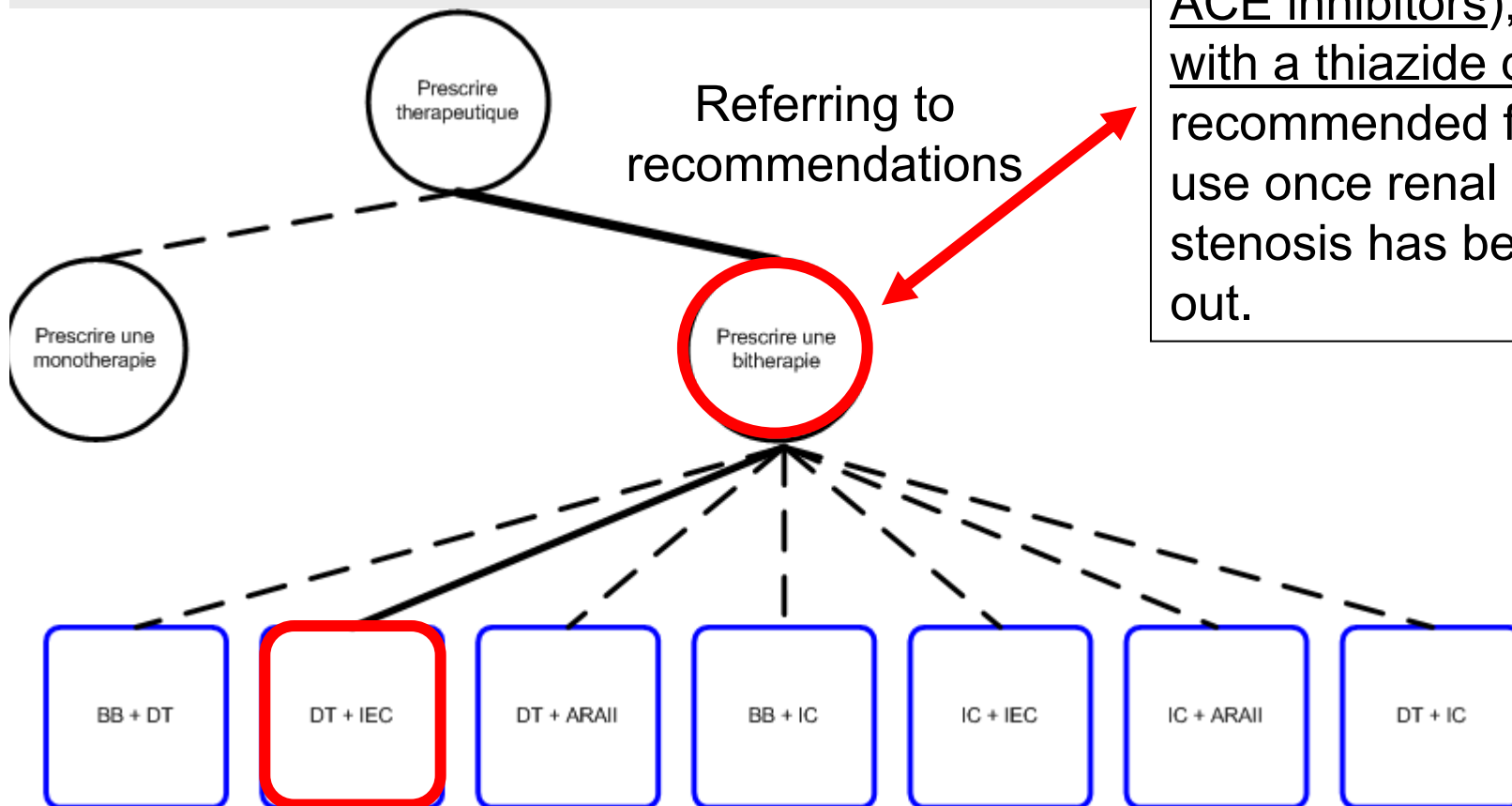
Une échographie prostatique par voie transrectale n'est pas recommandée lors du bilan initial de l'HBP symptomatique. Cet examen n'a aucune place dans le diagnostic, le bilan ni la surveillance d'une HBP.

Mapping HTN Traversal to the Guideline Document



- Recommendations interpreted as heuristics can be used to select the most appropriate alternative in the extraction of a solution task graph
- A key problem: how to interpret textual recommendations in terms of selection heuristics or grounded action costs?

Example of Heuristic Determination in Online Mode



Drugs acting on the renin-angiotensin system (i.e. ACE inhibitors), combined with a thiazide diuretic are recommended for first-line use once renal artery stenosis has been ruled out.

HTN Traversal with a Specific Clinical Situation (1/2)

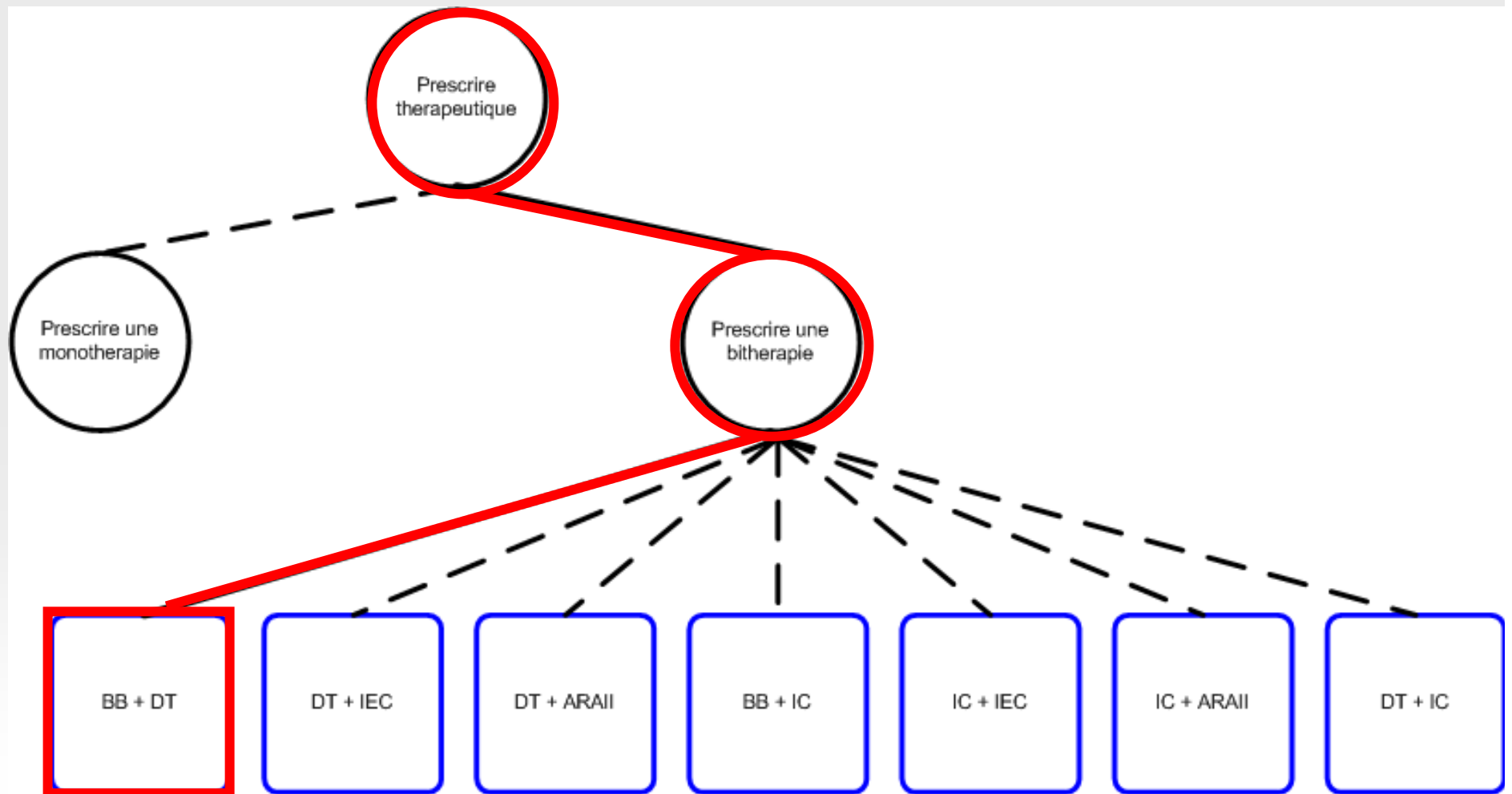
- Consider a patient with
 - a high level of cardiovascular risk and an antihypertensive monotherapy based on diuretics
- For whom an inefficient treatment

HTN Traversal with a Specific Clinical Situation (2/2)

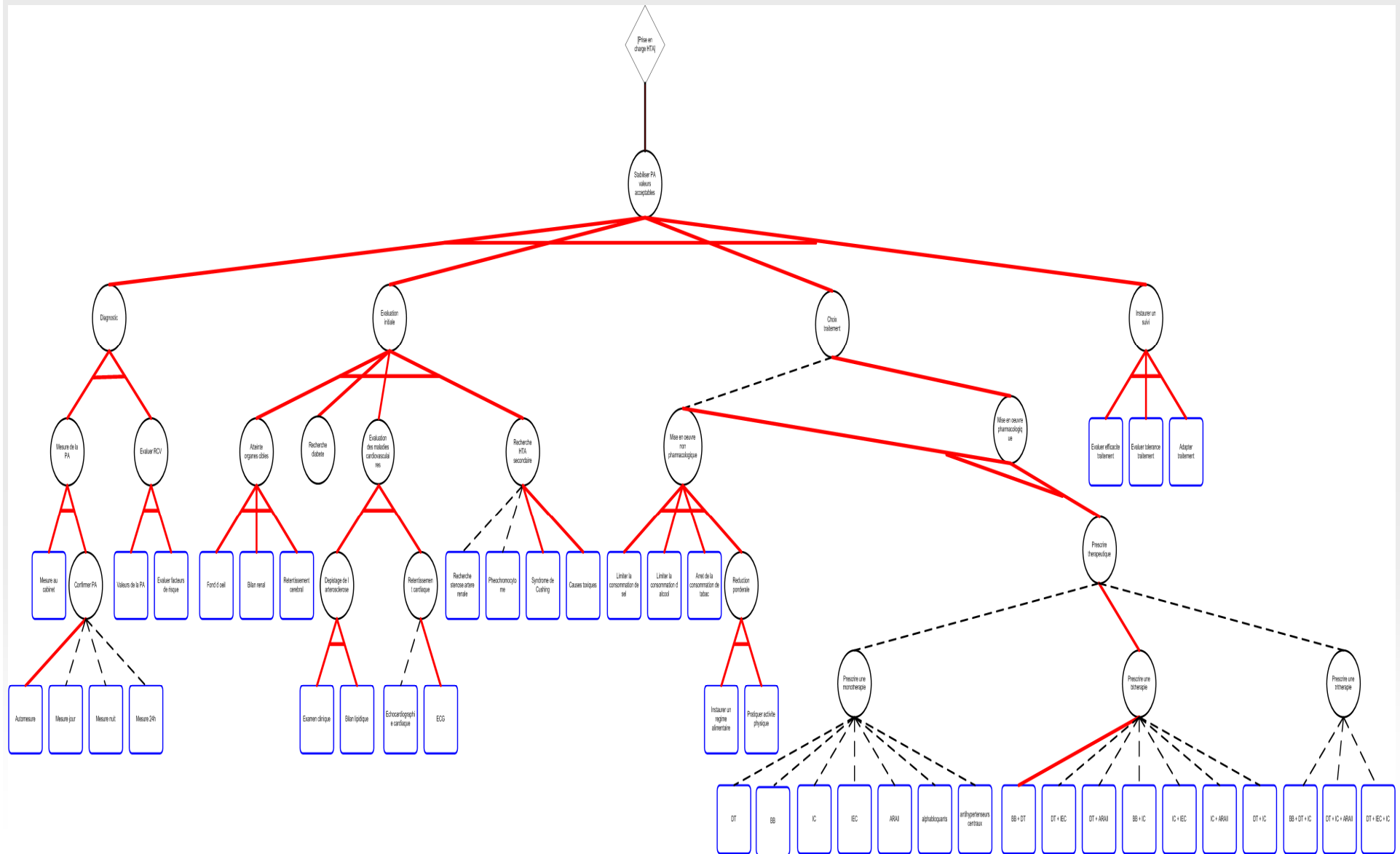
- The node “prescribe drug treatment” is selected in the HTN
- Three recommendations are highlighted in the clinical guidelines

- 1 • “If the BP target is not achieved with first-line therapy, a combination of two drugs may be started as second-line therapy after at least 4 weeks.”
- 2 • “However, it may be started early in patients with BP $\geq 180/110$ mm Hg regardless of the number of CVR factors; in patients with BP of 140-179/90-109 mm Hg and a high CVR.”
- 3 • “If the patient does not respond to the initial therapy after 4 weeks or experiences side effects, a drug from a different therapeutic class should be prescribed.”

Sub-graph Issues from the Instantiation of the Node



A solution Sub-graph



Conclusion

- The joint use of a knowledge model and a document model can be a new approach to study the completeness and consistency of clinical guidelines
- HTN provide a good formalization for protocols: textual recommendations can provide heuristics for solution extraction