CAHIER DU LAMSADE

Laboratoire d'Analyse et Modélisation de Systèmes pour l'Aide à la Décision
(Université Paris-Dauphine)
Unité de Recherche Associée au CNRS n° 825

ON AN APPROXIMATION MEASURE FOUNDED ON THE LINKS BETWEEN OPTIMIZATION AND POLYNOMIAL APPROXIMATION THEORY

CAHIER N° 116
septembre 1993

M. DEMANGE
V.Th. PASCHOS

received : March 1993.

CONTENTS

Sur une mesure d'approximation fondée sur les liaisons entre la théorie d'optimisation et celle de l'approximation polynomiale

Résumé

Il est bien connu qu'un bon moyen d'exprimer les problèmes de combinatoire est de les écrire comme des programmes linéaires en nombres entiers. Dans ce contexte, la théorie de l'approximation polynomiale de problèmes NP-complets semble incapable de cerner la nature des problèmes qu'elle traite en ce sens qu'elle ne permet pas de gérer du point de vue de l'approximation les différentes formulations, équivalentes au sens de l'optimisation, d'un même problème sous la forme de programme linéaire. En fait, la façon de formuler un problème influe considérablement, comme nous le mettons en évidence, les résultats d'approximation que nous pouvons obtenir. Nous discutons quelques anomalies de la théorie de l'approximation polynomiale qui semble ignorer la notion d'équivalence du point de vue de l'optimisation entre les formulations des problèmes. Nous commençons par définir formellement les notions de problèmes d'optimisation NP-complet et d'équivalence entre de tels problèmes incluant notamment quelques équivalences intuitives (comme, par exemple, la transformation d'un problème en composant l'objectif par une fonction affine) ou encore quelques équivalences évidentes entre problèmes de maximisation et de minimisation (comme par exemple l'équivalence entre entre *la couverture minimum d'ensembles* et *le stable maximum*), qui ne sont pas respectées dans le cadre de la théorie classique de l'approximation. Nous imposons comme principal axiome devant être vérifié par une mesure d'approximation le respect de cette équivalence et montrons qu'une mesure de la forme d'une fonction des deux variables - la valeur trouvée par l'algorithme - la valeur optimale - ne peut vérifier cet axiome. Nous introduisons alors un nouveau rapport d'approximation, comme fonction de trois variables, respectant cette notion d'équivalence et obtenu, sous le choix a priori de ces variables, par une approche axiomatique. Cette mesure a, entre autres, l'avantage d'uniformiser l'intervalle des valeurs possibles dans les cas de maximisation et de minimisation. Enfin, nous montrons quelques résultats d'approximation en utilisant ce rapport.

mots-clés: optimisation combinatoire, problème NP-complet, réduction polynomiale, approximation polynomiale

On an approximation measure founded on the links between optimization and polynomial approximation theory

Abstract

It is commonly known that a convenient way to express combinatorial problems is to describe them as integer linear programs. In this context, the theory of the polynomial approximation of NP-complete problems seems to be inadequate to capture the nature of the treated problems, as it is not able to capture in a global manner the equivalence—with respect to their approximation—of all the ways of expressing a given problem in terms of an integer program. In fact, as we show, the way a combinatorial problem is expressed, influences the approximation results one can obtain for it. We firstly define formally the notions of an optimization problem as well as the one of equivalence among such problems. This equivalence includes more particularly some intuitive equivalencies as the several ways of expressing an optimization problem (for example, by translating or affining the objective function) or yet some evident equivalencies between maximization and minimization problems (for example, the equivalence between minimum vertex cover and maximum independent set). We note here that the equivalencies mentioned before, are not covered by the conventional approximation theory. We impose as principal axiom on the approximation the respect of this equivalence and we prove that the approximation ratio defined as a two-variable function cannot verify this axiom. We then adopt a three-variable function as a new approximation ratio (already used by a number of researchers), which is coherent to the equivalence and, under the choice of the variables, the new ratio is introduced by an axiomatic approach. Finally, using the new ratio, we prove approximation results for a number of combinatorial problems.

keywords: combinatorial optimisation, NP-complete problem, polynomial time approximation

1 Introduction

Numerous researchers ([1,3,14]) working on the polynomial approximation of NP-complete problems, have been led in introducing a new approximation measure that takes into account the notion of the value of the *worst solution* (let us denoted this value by $\Omega(I)$; given an instance I of an optimization problem, the worst solution can be informally defined as the worst value solution, among the feasible ones for I). According to these researchers this new measure reflects better the structure of the problems treated. In fact, their main objection against the usual measure was that this measure creates an artificial dissymetry between minimization and maximization problems. Thus, in [20] has been shown that the problem of finding approximately optimal clusters can be solved in linear time under the maximization criterion (maximum cut) while it is P-complete under the minimization criterion (minimum clustering). But this phenomenon is in contradiction with the fact that these two problems are in fact "dual" and it occurs due to the non-stability of the classical approximation measure under linear transformations. In [1], a new measure has been proposed such that every ρ-approximation algorithm for the maximum cut problem is also a ρ-approximate algorithm for the minimum clustering problem (since in what follows we will not treat these two problems, we do not define them; their definitions can be found in [10]).
This is one of the numerous problems arising when the ratio $\frac{A(I)}{OPT(I)}$ (where $A(I)$ and $OPT(I)$ are the values of the approximate and the optimal solution for an instance I of an optimization problem) is adopted as measure of the quality of a heuristic for a NP-complete problem.
Another problem, very well known to the community of the researchers working on the polynomial approximability theory, is the one of the incompatibility (with respect to their approximations) between vertex cover and independent set. Given a graph $G = (V, E)$, a vertex cover is a subset $V' \subseteq V$ such that, for each edge $uv \in E$, at least one of u and v belongs to V' and the minimum vertex cover problem (VC) is to find a minimum size vertex cover; an independent set is a subset $V' \in V$ such that, for every pair (v_i, v_j) of vertices in V', the edge $v_i v_j$ does not belong to E, and the maximum independent set problem (IS) is to find a maximum size independent set. These two problems are linked by a simple relation, namely, given a graph, a vertex cover is the complement of an independent set with respect to the vertex set of the graph. The inconvenience of the traditionally adopted measure of the quality of an approximation algorithm is particularly apparent in the case of these two problems which are trivially equivalent (from an optimization point of view), in the sense that once the value of the solution of the former is given, a simple subtraction provides the value of the solution of the latter. However, this equivalence is not reflected on the approximation performance of the algorithms solving these two problems. For example, the maximal matching algorithm ([10,15]) is a 2-approximation algorithm for VC, while for IS, the approximation ratio of the same algorithm tends to infinity when the size of the instance also tends to infinity. Moreover, in [2] is shown that there is no constant-ratio approximation algorithm for IS unless P = NP.

However, in all the works cited above, the use of a measure taking into account the value of the worst solution was neither in the core of the problem treated, nor justified by itself,

except to the fact that it establishes a symmetry between minimization-maximization. On the other hand, we have to remark that the currently used ratio $\frac{A(I)}{OPT(I)}$ is also a priori selected, without any discussion or even "debate" on the properties such that a measure has to satisfy. But, the results obtained in polynomial approximation theory depend so largely on the adopted measure, that a debate on the choice of an approximation function seems to be very important. Thus, in the first part of this paper, we propose to "open" such a "debate" and to elaborate a formal answer. Proposition 1 for example, tries to justify the use of a third variable for the approximation function. Next, the use of the value of the worst solution (we will use the term *worst value* from now on) is motivated and, once the choice of variables has been performed, the form of this measure, that coincides with the one proposed originally in [1] and revisited in [14,3], appears as consequence of the axiom imposed on an approximation measure.
Up to now, this approximation function is only very punctually used for the evaluation of approximation algorithms, without its adoption to have a more general character. Only in [3] the authors use it in a systematic way, but only to study approximation preserving reductions and not to evaluate algorithms. So, in the second part of this paper and also in [7] we propose a systematic approach of the approximation theory using the new approximation ratio and we prove a variety of positive, negative and conditional results, also including in the latter approximation preserving reductions.
The starting point of this work consists in focusing on the main inconveniences of the current approximation measure. Its principal inconvenience is that it is not compatible with the optimization theory, the dissymetry between minimization and maximization being one of the multiple facets of this incompatibility. These remarks will lead us to the definition of a notion of equivalence between optimization problems that encompasses, for example, the links between maximum cut and minimum clustering or between VC and IS. This equivalence will motivate the choice of the principal axiom for the measure.

An optimization problem $\Pi_{1,0}$ can be written as:

$$\Pi_{1,0} = \underset{x \in S(I_{1,0})}{\text{opt}} \; f_{1,0}(\vec{x})$$

where $S(I_{1,0})$ is the set of the feasible solutions of the instance $I_{1,0}$ and the operator $\text{opt} \in \{\min, \max\}$.
For all pairs $(\lambda, \mu) \in I\!R^{+*} \times I\!R$, the problem:

$$\Pi_{\lambda,\mu} = \underset{x \in S(I_{1,0})}{\text{opt}} \; \lambda f_{1,0}(x) + \mu$$

where $I_{1,0}$ is always any instance of $\Pi_{1,0}$, is equivalent to $\Pi_{1,0}$ (in the sense that $\Pi_{1,0}$ and $\Pi_{\lambda,\mu}$ have the same set of optimal solutions.
The transformation from $\Pi_{1,0}$ to $\Pi_{\lambda,\mu}$ is so natural that we can consider all problems $\Pi_{\lambda,\mu}$, $(\lambda, \mu) \in I\!R^{+*} \times I\!R$ as equivalent versions of the same problem Π.
Given an approximation algorithm $\mathcal{A}$ that solves (approximately) a NP-complete problem Π, a usual measure of its approximation performance is the approximation ratio of $\mathcal{A}$ defined as $R_{\mathcal{A}}(I_{1,0}) = \frac{A(I_{1,0})}{OPT(I_{1,0})}$, where $OPT(I_{1,0})$ is the value of $\Pi_{1,0}$ and $A(I_{1,0})$ is the value of the solution found by algorithm $\mathcal{A}$ when applied to the instance $I_{1,0}$.

Of course, given an instance $I_{\lambda,\mu}$ of $\Pi_{\lambda,\mu}$, $OPT(I_{\lambda,\mu}) = \lambda OPT(I_{1,0}) + \mu$, the value of the solution provided by $\mathcal{A}$ becomes $\lambda A(I_{1,0}) + \mu$, and consequently, the approximation ratio of $\mathcal{A}$ for the new (equivalent) problem is now defined as $R_{\mathcal{A}}(I_{\lambda,\mu}) = \frac{\lambda A(I_{1,0})+\mu}{\lambda OPT(I_{1,0})+\mu}$.
Obviously, from a mathematical point of view, the two ratios $R_{\mathcal{A}}(I_{\lambda,\mu})$ and $R_{\mathcal{A}}(I_{1,0})$ are equally interesting and natural (the example of IS and VC is meaningful here). The same remark occurs for both alternatives we use to express the problem Π. We show below that the choice of λ and μ affects considerably the approximation results one can obtain.

For a given problem $\Pi_{1,0}$, we suppose that $A(I_{1,0})$ and $OPT(I_{1,0})$ are functions of the size of the problem which are bounded when the size is bounded. Obviously, all the known algorithms verify this hypothesis.
We have then, for $I_{1,0}$ of size less than n, $\lim_{n\to\infty} \frac{A(I_{1,0})+\mu}{OPT(I_{1,0})+\mu} = 1$. Therefore, for a μ sufficiently big, it holds that $\forall \Pi_{1,0}, \forall K \in I\!N, \forall \epsilon, \forall \mathcal{A}, \exists \mu$ such that $R_{\mathcal{A}}(I_{1,\mu}) \leq 1+\epsilon$ for all $I_{1,\mu}$ of size less than K.

For the problems admitting an algorithm which approximates their optimal solutions within a constant ρ, we can obtain a stronger result by treating λ and μ as universal constants (not depending on the instance).
Let the problem $\Pi_{1,0}$ and the algorithm $\mathcal{A}$ be such that $\forall I_{1,0}$ $\frac{A(I_{1,0})}{OPT(I_{1,0})} \leq \rho$ for a fixed constant ρ. Then for every ϵ, there exists $\lambda_\epsilon, \mu_\epsilon$ such that, for all $I_{\Pi_{1,0}}$, $1 \leq R_{\mathcal{A}}(I_{\lambda_\epsilon,\mu_\epsilon}) \leq 1+\epsilon$.
In fact, $R_{\mathcal{A}}(I_{\lambda,\mu}) \leq \frac{\lambda\rho OPT(I_{\Pi_{1,0}})+\mu}{\lambda OPT(I_{\Pi_{1,0}})+\mu} = \rho - (\rho-1)\frac{\mu}{\lambda OPT(I_{1,0})+\mu} = \rho - f(OPT(I_{1,0}))$. Without loss of generality, let us suppose that $\Pi_{1,0}$ is a minimization problem, thus $\rho > 1$, $R_{\mathcal{A}}(I_{\lambda,\mu}) \geq 1$. Then $f(OPT(I_{1,0}))$ is a decreasing function of $OPT(I_{1,0})$ whenever $\mu \leq 0$ (for a maximization problem, we would have taken $\mu \geq 0$).
Let $\sigma = \min_{I_{1,0}} OPT(I_{1,0})$, $\sigma \in I\!N^+$. Then, for all $I_{\lambda,\mu}$, $(\lambda,\mu) \in I\!R^{+*} \times I\!R$, we have $R_{\mathcal{A}}(I_{\lambda,\mu}) \leq \rho - (\rho-1)\frac{\mu}{\lambda\sigma+\mu}$. Note that if $\sigma = 0$, $R_{\mathcal{A}}(I_{\lambda,\mu}) = 1$ for all $I_{\Pi_{\lambda,\mu}}$, $\mu \neq 0$.
Hence $\lim_{\mu\to\infty} \rho - (\rho-1)\frac{\mu}{\lambda\sigma+\mu} = 1^+$ and consequently, by choosing μ arbitrarily small, we get $1 \leq R_{\mathcal{A}}(I_{\lambda,\mu}) \leq 1+\epsilon$.
Therefore, algorithms as for example the one of maximal matching for vertex cover, can have approximation ratios arbitrarily close to 1 if we make a judicious choice of λ and μ. Hence, given a problem Π and all equivalent forms $\Pi_{\lambda,\mu}$, the choice of λ and μ in the objective function influences considerably the value of the approximation ratio of an algorithm supposed to solve Π (and also all $\Pi_{\lambda,\mu}$). This inconvenience seems to us quite embarrassing, as it isolates complexity theory from combinatorial optimization theory, both theories been complementary for almost all researchers.
For the stability of any approximation measure with respect to $\Pi_{1,0}$ and $\Pi_{\lambda,\mu}$, we prove the following.

Proposition 1. *Let Π be an NP-complete problem and let $\Pi_{\lambda,\mu}$, $(\lambda,\mu) \in I\!R^{+*} \times I\!R$ be the versions of Π arising from a strictly increasing affine transformation of its objective function. Let also R be an approximation measure of an approximation algorithm for Π and let $R_{\lambda,\mu}, R_{1,0}$ be its values for $\Pi_{\lambda,\mu}, \Pi_{1,0}$, respectively. Then, the measure R for which $R_{\lambda,\mu} = R_{1,0}$, $\forall(\lambda,\mu) \in I\!R^{+*} \times I\!R$ holds, cannot be expressed as a 2-variable function $R(A(I), OPT(I))$.*

Proof: Indeed, such a function should satisfy the following property:

$$\forall (x,y) \in I\!R^2, \forall (\lambda,\mu) \in I\!R^{+*} \times I\!R : \quad R(\lambda x + \mu, \lambda y + \mu) = R(x,y) \ .$$

If we denote by $\mathcal{E}_{x,y}$ the set $\{(\lambda x + \mu, \lambda y + \mu) : (\lambda,\mu) \in I\!R^{+*} \times I\!R\}$, we have obviously:

$$\begin{array}{lll} x = y & \Longrightarrow & \mathcal{E}_{x,y} = \{x = y\} \qquad \text{(i)} \\ x > y & \Longrightarrow & \mathcal{E}_{x,y} = \{x > y\} \qquad \text{(ii)} \\ x < y & \Longrightarrow & \mathcal{E}_{x,y} = \{x < y\} \qquad \text{(iii)} \end{array}$$

In fact, $\mathcal{E}_{x,y} = \bigcup_{\lambda>0} \{(\lambda x + \mu, \lambda y + \mu)\mu \in I\!R\} = \bigcup_{\lambda>0} D_\lambda$, where D_λ denotes the straight line parallel to $\{x = y\}$, passing through the point $(\lambda x, \lambda y)$. $\mathcal{E}_{x,y}$ appears to be the set of the translations of the straight line $x = y$ with vectors $k \cdot \binom{x}{y}, k > 0$.
In case (i), these vectors are all to the direction of D_λ; in cases (ii) and (iii), the vectors of translations are not to the direction of D_λ, thus in both cases, $\mathcal{E}_{x,y}$ is a half-plane.
Consequently, R should be constant in all the three sets $\{x = y\}$, $\{x > y\}$, $\{x < y\}$. Such an approximation ratio would just detect if $A(I) = OPT(I)$, $A(I) > OPT(I)$, or $A(I) < OPT(I)$, respectively, and nothing more. But such a ratio does not give any meaningful information at all. ∎

Another problem someone could mention here, is that the traditional definition of the approximation ratio gives too little information on the real quality of an approximation algorithm. In fact, let us take the case of an algorithm solving approximately problem Π within a constant error. Most of the known algorithms which they have this nice property are greedy algorithms that, in many cases, provide the worst value feasible solution. The maximal matching algorithm for VC is such an algorithm. We think that those algorithms, even if they are considered as efficient, are not outstanding. Their "good" behaviour is rather due to some "nice" properties of the optimal solution than to an efficient conception (and operation) of the algorithm itself. Always in the case of VC, it is commonly known that almost every graph admits a perfect matching ([5]) and also that, in every graph, the covering number is greater or equal to the cardinality of a maximum matching ([4]). Thus, in many cases, a solution containing all the vertices of the graph has cardinality not greater than twice the cardinality of the optimal one. On the other hand, one can remark that almost every graph has maximum independent set size $O(\log n)$ ([5]) and thus a VC solution of cardinality equal to the order of the graph has approximation ratio asymptotically equal to one. But isn't it a default of the usual measure the fact that we cannot affirm this result at least for large classes of VC instances?
It seems to us that hoping to optimize (minimize or maximize) a solution is synonymous to trying to find a solution as far as possible from the worst value one and therefore the approximation ratio in use is not the best measure for this taking away.

All the previous reasons make us to consider that the conciliation of the quality measure someone uses to estimate the approximation performance of an algorithm and the intuitive ideas on the equivalence between optimization problems is an important problem.
In the following, we define the problems treated by the approximation theory, which are the optimization versions of the decision problems defined formally in [10]. We define a

notion of equivalence between such problems, equivalence that respects some commonly accepted (from an optimization point of view) affinities between some problems, like instance the vertex covering and the independent set. Finally, the formalization of our notion of equivalence, allows us to propose a new approximation measure which follows this equivalence.
The present version of the paper is the detailed version of [6].

2 A notion of equivalence and some of its impacts

Usually, a problem is formally defined in a decision version, i.e. in terms of a question about the existence of a solution satisfying some required constraints. An instance of a problem can be defined either by the specification of particular values of problem parameters, or by the specification of the question itself (by attributing particular values to the parameters of the constraints). In this paper, we adopt this second approach.
But there exist problems consisting in finding either a solution (if such solution exists, as for example in the case of satisfiability or of Hamiltonicity), or the best (minimum or maximum) value solution between the existing ones; these last problems also admit a decision version.
Commonly, the theory of the approximation of NP-complete problems focuses on "optimization problems" in the sense that we need to attribute to each solution a value characterizing the quality of this solution; consequently, we are interested in "optimization problems" (having NP-complete decision version), these problems being defined in definition 1.

Definition 1. Let opt $\in \{\max, \min\}$, let $E_1, \ldots, E_p$ be a finite number of closed subsets of $\mathbb{R}$. An *elementary affine optimization problem* is a sextuple $(n, t, s, v, C, (A_i)_{i \in \{E_1, \ldots, E_p\}})$, where $n \in \mathbb{N}$ is the size of the elementary problem, t is a dimension and s is a number of constraints both (t and s) polynomial in n, v is an affine function on $\mathbb{R}^t$, C is a polytope of $\mathbb{R}^t$ defined by the s constraints and finally $A_i \in \{E_1, \ldots, E_p\}, i \in \{1, \ldots, t\}$ (for the moment, we restrict ourselves to problems for which there is always a solution (i.e. E_i closed and C compact).
We call such an elementary problem an *instance* and we write it in the form

$$\begin{cases} \text{opt} \quad v(\vec{x}) & \\ \qquad\quad \vec{x} & \in \; C \\ \qquad\quad x_i & \in \; A_i, \, i \in \{1, \ldots t\} \end{cases}$$

The *size of an instance* is n.
An *optimization problem*, is a set of instances such that the set of questions "$\forall M$ is there any $\vec{x} \in C$ such that $v(\vec{x}) \theta M$?" (where θ equals $\leq$ (resp. $\geq$) if opt equals max (resp. min)) is exactly the set of instances of a problem in the usual sense[1]. ∎

[1] In general, to express the fact that all instances of a given problem have the same structure, arguments of language theory are used ([10,13]); we have given this definition in order to avoid to repeat such a discussion.

An *algorithm* is defined, as usual, as a finite sequence of elementary steps consisting in selecting an element in E_i and performing all the arithmetical or logical operations on it. The complexity of an algorithm is defined as the number of elementary steps needed to complete the solution.

Definition 2. Two problems Π_1, Π_2 are *equivalent* if
(a) there exists an algorithm $\propto_{12}$ polynomial in n_1, which from every instance I_1 of Π_1 (n_1 is the size of I_1) constructs an instance I_2 of Π_2 (we call algorithm $\propto_{12}$ a reduction from Π_1 to Π_2), where n_2 (the size of I_2) is a polynomial of n_1; there exists a polynomial time algorithm ψ_{21} such that for each instance I_1 and $I_2 = \propto_{12} (I_1)$ and $\forall x \in C_2$, $\psi_{21}(x) \in C_1$, where C_1, C_2 are the constraints corresponding to Π_1, Π_2 (the expression ψ_{21} is a polynomial time algorithm which means that it works in time polynomial in the size of its variable x, which is by definition the size of the corresponding instance); there exists a bijective affine function $f_{21} : I\!R \mapsto I\!R$ preserving the order of the solutions from their value point of view (f_{21} increases if Π_1, Π_2 are both maximization or minimization problems and f_{21} decreases in the opposite case, i.e. one of Π_1, Π_2 is maximization and the other one is minimization problem), and for which holds: $(\forall x \in C_2)\ v_1(\psi_{21}(x)) = f_{21}(v_2(x))$;
(b) there is a reduction $\propto_{21}$, polynomial in n_2, which constructs an instance I_1 of Π_1 from every instance I_2 of Π_2, where n_1 is a polynomial of n_2; there exists a polynomial time algorithm ψ_{12} and a bijective affine function f_{12} for each instance I_2 and $I_1 = \propto_{21} (I_2)$ satisfying **(a)** via the interchanging of indices 1 and 2;
(c) $\propto_{12} \circ \propto_{21} = Id$, $\propto_{21} \circ \propto_{12} = Id$; for every instance I_1, $(\psi_{12} \circ \psi_{21})|_{C_1} = Id_{C_1}$ and for every instance I_2,$(\psi_{21} \circ \psi_{12})|_{C_2} = Id_{C_2}$. ∎

We note here that $\psi_{i,j} (i, j \in \{1, 2\}, i \neq j)$ map the best solution of Π_i to the best solution of P_j.
Also, when two problems are equivalent, ψ_{12}, ψ_{21} associate bijectively C_1, C_2 in such a way that the distribution functions of I_1, I_2 are identical in the sense that

$$\forall\ k \in\ I\!R, \Pr[v_1(x_1) =\ k] = \Pr[v_2(x_2) = f_{12}(k)].$$

This property can be interpreted in terms of *structure*, notion introduced in [3]. There, the structure of an instance I is defined as the list $(a_1, \cdots, a_c)$ where $c = |OPT(I) - \Omega(I)|$ and $a_i = |\{x \in C_I, |v(x) - \Omega(I)| = i\}|$. This definition as well as the one of a *convex problem* (always in [3]) have the drawback to be incompatible with the homothetic transformations. Let us, for example, consider IS whose an instance can be formulated as

$$\text{IS} = \begin{cases} \max & \vec{1}_{|V|} \cdot \vec{x} \\ & A \cdot \vec{x} \ \leq\ \vec{1}_{|E|} \\ & \vec{x} \ \in\ \{0,1\}^{|V|} \end{cases}$$

where A is the edge-vertex incidence matrix of the graph representing an instance of IS. This problem is convex in the sense of [3], and this means that all the integer values in $[OPT(I), \Omega(I)]$ can be attained. On the other hand, the problem

$$\text{IS}' = \begin{cases} \max & \vec{2}_{|V|} \cdot \vec{x} \\ & A \cdot \vec{x} \ \leq\ \vec{1}_{|E|} \\ & \vec{x} \ \in\ \{0,1\}^{|V|} \end{cases}$$

where $\vec{2}$ is the vector with all its components equal to 2, is not convex as the odd values are never attained. For this reason IS$'$ has not the same structure as IS (always in the sense of [3]).
In our context on the contrary, it is natural for both problems to have the same structure. We define thus the *structure* of an instance (we prefer to call it by *sharing out of the values* of an instance) as the equivalence class defined by the following equivalence relation: $(a_0, \cdots, a_n) \equiv (b_0, \cdots, b_m) \Longleftrightarrow (n = m) \ \vee \ \exists \lambda \in I\!R^*, \forall i \in \ \{0, \cdots, n\} \ a_i = \lambda \cdot b_i$ where $(a_0, \cdots, a_n)$ and $(b_0, \cdots, b_m)$ are structures in the sense of [3].
It is easy to see that, under this definition, the equivalence preserves values sharing out. Moreover, we note that all the structure preserving reductions described in [3] remain structure preserving for our definition.

Finally, we should note that the main difference between the equivalence of definition 2 and the strong equivalence defined in [11] is that the authors of [11] consider $f_{21} = Id$. But this consideration does not take into account the notion of the affine transformation, notion central for us. More particularly, the definition of [11] does not allow the equivalence between a minimization and a maximization problem, for example between VC and IS.

The form of a new approximation ratio

Let Π be an NP-complete optimization problem and $\mathcal{A}$ an approximation algorithm for Π. We choose, as in the classical approximation theory, to associate a real number ρ to the pair $(\Pi, \mathcal{A})$; this number is called approximation measure of $\mathcal{A}$ for Π. In fact, as usual, ρ will be defined to reflect the behaviour of $\mathcal{A}$ on each instance of Π. Thus, we start by associating to every instance I of Π a number $f(I, \mathcal{A})$; by convention $f \in [0, 1]$. Here resides the first difference with respect to the classical approach which differentiates the interval of the possible values of f depending on the nature (maximization or minimization) of the objective. Once this convention been made, the first property that seems natural to be imposed to f is that it has to increase when $A(I)$ is improved (in the sense of the objective), by taking the value 1 when $\mathcal{A}$ finds the optimal solution of I.
Starting from the local measure f let us now define a global one. Recall that we are placed in the frame of the *worst case complexity*, in the sense that the approximation results we try to obtain provide a guarantee over all instances of a given problem for the considered algorithm. It is natural then, that for a given problem, a global measure represents the worst behaviour of an algorithm on the different instances of the problem. As for $f(\mathcal{A}, I)$ the nearer to 1 the better, we define:

$$\rho(\Pi, \mathcal{A}) = \inf_{I \in \mathcal{I}(\Pi)} (f(\mathcal{A}, I))$$

where, $\mathcal{I}(\Pi)$ the set of instances of Π. In the following, for purposes of clarity, when no confusion arises we will write ρ instead of $\rho(\Pi, \mathcal{A})$.
Let us note that the defined measure verifies the following natural "growth property":

Property. *Whenever there exist algorithms $\mathcal{A}, \mathcal{A}'$, for Π, if $\mathcal{A}$ gives better solution values for an instance I of Π, then the value of the approximation measure for $\mathcal{A}$ is better (nearer to 1) for I than the one of $\mathcal{A}'$.*

This is a very natural and general property verified by every approximation measure which mainly implies the monotonicity for an approximation function.
On the other hand, in order to answer to the problematics we are imposed we define a second property taken as an axiom for the new approximation approach.
When speaking of axiom, we mean whatever is logic and coherent to require from an approximation measure under the light of the thoughts and reserves which have been expressed above. Let us recall that the essential criticism against the usual measure was that it does not respect the notion of equivalence between NP-complete optimization problems. By having introduced such a notion of equivalence in definition 2, we propose the following principal axiom:

Axiom. *If there are two equivalent problems, then every algorithm solving the former allows to find (modulo ψ) an algorithm solving the latter within the same value of the approximation measure.*

This, in terms of [22], signifies that algorithms $\propto$ of definition 2 constitute continuous (approximation preserving) reductions and also that the algorithms ψ are the two $1,0$-bounded, corresponding functions. In fact in the frame of the new approximation ratio, one can see the equivalence as the 0-level of the continuous reductions.
We recall also that, in proposition 1, we have proved that a measure seen as a 2-variable function $R(A(I), OPT(I))$ cannot respect the above axiom because already it cannot respect it when f_{21} is increasing.
However, both variables $A(I), OPT(I)$ seem to be very meaningful and natural because, from our point of view, the predominating notion for estimating an algorithm is the value of the solution it provides. On the other hand, we have already discussed the fact that an approximation algorithm provides many times the worst value solution (Ω) for some instances of a problem.
In fact, given an optimization problem

$$\Pi = \begin{cases} \text{opt} \quad v(\vec{x}) & \\ \qquad\quad \vec{x} & \in C \\ \qquad\quad x_i & \in A_i,\ \forall i \in \{1, \ldots, t\} \end{cases}$$

we define Ω to be the value of the problem

$$\Pi' = \begin{cases} \overline{\text{opt}} \quad v(\vec{x}) & \\ \qquad\quad \vec{x} & \in C \\ \qquad\quad x_i & \in A_i,\ \forall i \in \{1, \ldots, t\} \end{cases}$$

where $\overline{\text{opt}} = \min(\max)$ if $\text{opt} = \max(\min)$.
In [7], we show how the formulation of an optimization problem contains implicitly the

notion of Ω as it contains explicitly the one of the optimum of the objective function.
The arguments presented up to now make us to think that the parameter Ω is a consistent variable that has to be considered for the adoption of a new approximation function.
Therefore, it seems natural to introduce as an approximation ratio, a 3-variable function $R(A(I), OPT(I), \Omega(I))$.
In any case this notion of worst value is worthy of many remarks. First, let us take the work in [3]. There, the authors call *trivial value* what we call here trivial value. However, it should be better not to conclude that it is always easy to calculate it; if there are many cases where the computation of the worst value is easy, there are many problems for which the computation of the worst value is as difficult as the computation of the optimal one, i.e. NP-complete. Let us for example consider the case of the minimum maximal independent set in a graph. This problem is exactly the one of searching for a minimum dominating set constituting simultaneously an independent set and it is NP-complete ([10]). The worst value solution for this problem is the independence number of a graph, and the problem to be solved in order to find it, is exactly the IS. In [8], a broad class of examples is given, and it is shown that the notion of worst value is crucial for the definition of a problem, and moreover that in many cases, changing the worst value solution entails the changing of the optimization problem itself.

Let us consider two equivalent problems Π_1, Π_2 (definition 2).
We suppose also that we have an approximation algorithm $\mathcal{A}_1$ solving all instances of Π_1.
We consider the following algorithm $\mathcal{A}_2$ for Π_2: $\forall I_2, \mathcal{A}_2(I_2) = \psi_{12}(\mathcal{A}_1(\propto_{21} (I_2))$.
If $\lambda \in I\!R^*, \mu \in I\!R$ are the two numbers such that $\forall x \in I\!R, f_{12}(x) = \lambda x + \mu$, we have respectively: $OPT(I_2) = \lambda OPT(\propto_{21} (I_2)) + \mu$, $A_2(I_2) = \lambda A_1(\propto_{21} (I_2)) + \mu$, $\Omega(I_2) = \lambda\Omega(\propto_{21} (I_2)) + \mu$, where A_1 (A_2) denotes the value of the solution provided by the algorithm $\mathcal{A}_1$ ($\mathcal{A}_2$), respectively.
We denote by $\mathcal{E} = \{(x, y, z) \in I\!R^3 : x \neq z \text{ and } y \text{ between } x \text{ and } z\}$, the set of all the possible triplets for $(OPT(I), A(I), \Omega(I))$ (in fact, the case $x = z$ corresponds to the case where $OPT(I) = \Omega(I) = A(I)$, which is trivial).
Then (following the imposed axiom), the 3-variable function R, has to satisfy $R(\lambda x + \mu, \lambda y + \mu, \lambda z + \mu) = R(x, y, z)$, $\forall (x, y, z) \in \mathcal{E}, \forall (\lambda, \mu) \in I\!R^* \times I\!R$.
Let us denote by $\mathcal{F}$ the set $\{(\lambda x_0 + \mu, \lambda y_0 + \mu, \lambda z_0 + \mu), (\lambda, \mu) \in I\!R^* \times I\!R\}$ and by $\mathcal{G}$ the set $\{(x, y, z) \in \mathcal{E} : \frac{y-z}{x-z} = \frac{y_0-z_0}{x_0-z_0}\}$; it is obvious that $\mathcal{F} \subset \mathcal{G}$. We prove that $\forall (x_0, y_0, z_0) \in \mathcal{E}$, $\mathcal{F} = \mathcal{G}$.
Consider a triplet $(x, y, z) \in \mathcal{G}$ and, define $\lambda = \frac{x-z}{x_0-z_0}$ (λ exists because $x_0 \neq z_0$) and $\mu = y - \lambda y_0$.
We have: $z = y - \lambda y_0 + \lambda z_0 = \lambda z_0 + \mu$, $x = z - \lambda z_0 + \lambda x_0 = \lambda x_0 + \mu$, hence $\mathcal{G} \subset \mathcal{F}$.
Consequently, R is constant on each hyperplane $\frac{y-z}{x-z} = constant$ and thus, there exists a function $f : I\!R \mapsto I\!R$ which satisfies $R(x, y, z) = f\left(\frac{y-z}{x-z}\right)$, $\forall (x, y, z) \in \mathcal{E}$.
Let us consider two algorithms $\mathcal{A}$, $\mathcal{A}'$ for a given problem, the first algorithm being better (more efficient) than the second one. This means that for every I, $\frac{A(I)-\Omega(I)}{OPT(I)-\Omega(I)} \geq \frac{A'(I)-\Omega(I)}{OPT(I)-\Omega(I)}$. By the property imposed on the ratio we have then $|1 - f\left(\frac{A(I)-\Omega(I)}{OPT(I)-\Omega(I)}\right)| \leq |1 - f\left(\frac{A'(I)-\Omega(I)}{OPT(I)-\Omega(I)}\right)|$.
If we suppose for instance that the approximation ratio is less than 1 (and consequently

maximum in the best case), then $f : [0,1] \mapsto [0,1]$ and it is decreasing.

We choose the simpler function f satisfying the previously imposed axiom and property the identity function.
Consequently, we propose as an approximation measure the ratio

$$\frac{A(I) - \Omega(I)}{OPT(I) - \Omega(I)}.$$

If $A(I) = OPT(I)$, then the ratio is equal to 1, and if $A(I) = \Omega(I)$, then it is equal to 0. If $OPT(I) = \Omega(I)$, then the ratio is undefined but this case is trivial for the approximation. A possible interpretation of this ratio is that it expresses the position of $A(I)$ into the interval of its possible values.

The proposed ratio, obviously satisfies the imposed axiom, i.e. if two problems Π_1, Π_2 are equivalent, then for every algorithm $\mathcal{A}_1$ solving the first problem, we can find in polynomial time—as we have explained above—an algorithm $\mathcal{A}_2$ having the same ratio for every instance of the second problem.
Here, we have to note that any ratio obtained by a circular permutation of the three parameters $A(I), OPT(I), \Omega(I)$ (for example $\frac{OPT(I)-A(I)}{A(I)-\Omega(I)}$) satisfies both the axiom and the property. The only thing that changes eventually, if we adopt such a ratio is the interval of variation of the ratio's values.

In fact, someone can note that the axiom and the property imposed on the ratio are satisfied even if we use a more restricted version of equivalence called weak equivalence.

Definition 3. Two problems Π_1, Π_2 are weakly equivalent if
(a) there exists an algorithm $\propto_{12}$ polynomial in n_1, which from every instance I_1 of Π_1 constructs an instance I_2 of Π_2, where n_2 is a polynomial of n_1; there exists a polynomial time algorithm ψ_{21} such that for each instance I_1 and $I_2 = \propto_{12} (I_1)$ and $\forall x \in C_2$, $\psi_{21}(x) \in C_1$; there exists a bijective affine function $f_{21} : \mathbb{R} \mapsto \mathbb{R}$ such that for each instance I_1 and $I_2 = \propto_{12} (I_1)$ and $\forall x \in C_2$, $v_1(\psi_{21}(x)) = f_{21}(v_2(x))$, $f_{21}(OPT(I_2)) = OPT(I_1)$, $f_{21}(\Omega(I_2)) = \Omega(I_1)$, where $\Omega(I_i)$ denotes the value of the worst solution of the instance I_i of the problem Π_i;
(b) there is a reduction $\propto_{21}$, polynomial in n_2, which constructs an instance I_1 of Π_1 from every instance I_2 of Π_2, where n_1 is a polynomial of n_2; there exists a polynomial time algorithm ψ_{12} and a bijective affine function f_{12} for each instance I_2 and $I_1 = \propto_{21} (I_2)$ satisfying **(a)** via the interchanging of indices 1 and 2. ∎

In order to obtain positive results using the adopted measure, one has to deviate, more or less, from the way he/she was used to think up to now, mainly in the case of minimization problems. Major notion conceptually included in this new ratio is the taking away for the worst value feasible solution of an instance. This taking away is quite explicit for the most of the maximization problems where, usually, we start from an empty set constituting a feasible solution (the worst value one) and we try to augment it while it is possible (taking

away from the worst case). This process constitutes a strategy applied even by the greedy algorithms. On the contrary, for minimization problems the taking away is quite rare and a strategy where we start from the worst value solution by trying to reduce it is either non efficient or it leads quickly to non-feasible solutions. For example, in the case of VC if we start from the worst value solution, (the vertex set of the graph) and we try to reduce its size by obtaining a smaller one, we are unable to find bounds for the obtained cardinality. We think that an interesting idea for overcoming this difficulty is to use the advantages definition 2 provides us. In [7] we give examples of how one the defined equivalence can be used, in order to produce positive results for minimization problems.

Let us revisit for a while the notion of the sharing out of the values of an instance. If this sharing out is known (but this is impossible unless P = NP), then one can interpret a value χ of the ratio as the fact that the solution found lies among the $(1-\chi) \times 100$ percent of the better solutions. This measure can be considered by many people as the ideal approximation measure. Unfortunately, positive results using such a measure are impossible.
In fact, the possibility to prove results, and in particular positive ones, validates a posteriori every proposition of an approximation ratio. Our approach seems to guarantee a consistent ratio, even if other ratios defined in a different way could contain even a greater amount of information. On the other hand, the present approach allows us to produce results as we show in the sections that follow.
But, the scientific richness and wealth of this theory does not consists only of its capacity to contain positive results but also to allow negative or conditional ones. In particular a theory where only positive results are proved is dull and without a real scientific content. On the other hand a theory only allowing negative results, being too restrictive and inefficient. In what follows we show that the adopted theory is rich enough to allow all kinds of results dealing very well with the NP-completeness theory, which contains problems of various degrees of difficulty concerning their approximation.

3 Approximation results obtained using the new approximation ratio

3.1 The knapsack problem

We treat now the Bivalued Knapsack Problem (KP). This problem is usually defined in its decision version as follows ([10]).

Bivalued Knapsack Problem (KP). Given a finite set $I = \{1, \ldots, n\}$, with sizes $a_i \in I\!N^*$, a value $c_i \in I\!N^*$ and two positive integers b and d, is there a subset $I' \subseteq I$ such that $\sum_{i \in I'} a_i \leq b$ and such that $\sum_{i \in I'} c_i \geq d$?

Of course when speaking of approximation, we have to consider the optimization version of the corresponding problem. In this case two optimization versions appear naturally: the maximization and the minimization ones. Both versions are respectively defined as

follows:

$$\text{KM} = \begin{cases} \max & \vec{c}\cdot\vec{x} \\ & \vec{a}\cdot\vec{x} \leq b \\ & x_i \in \{0,1\},\ \forall i \in I \end{cases}$$

$$\text{Km} = \begin{cases} \min & \vec{a}\cdot\vec{x} \\ & \vec{c}\cdot\vec{x} \geq d \\ & x_i \in \{0,1\},\ \forall i \in I \end{cases}$$

For KM, Ibarra and Kim ([12], see also [21]) present a polynomial time approximation schema (PTAS) by using the traditional approximation ratio. We treat in this paper the following generalizations of KM and Km called RKM, RKm, where the components of the vectors $\vec{a}$ and $\vec{c}$ as well as b and d are real numbers.
For these problems we propose a PTAS valid for the approximation ratio defined in the previous section. Obviously, by the definition of our ratio, in the case of RKM where the components of $\vec{a}$ and $\vec{c}$ and the constant b are all positive numbers (PRKM), this schema works also for the traditional ratio. We note that this case is a generalisation of the results of [12].
We shall see that the adopted approximation ratio plays a crucial role as it allows such generalizations. We will see also that, following the notion of equivalence discussed above, KM and Km are equivalent as well as RKM, RKm and PRKM.
We discuss after how the PTAS for KM and Km gives also a pseudo-polynomial algorithm for subset sum problem (S) and also how to generalize S to preserve the pseudo-polynomiality. This discussion, seen in the case of the PTAS for RKM and RKm, permits to understand why the subset product problem (SP) is NP-complete in the strong sense. S and SP are defined in their decision versions as follows.

Subset Sum Problem (S). Given a finite set A with a size $s(a) \in Z^+$ for each $a \in A$ and a positive integer B, is there any subset $A' \subseteq A$ such that the sum of the sizes of the elements in A' is exactly B?

Subset Product Problem (SP). Given a finite set A with a size $s(a) \in Z^+$ for each $a \in A$ and a positive integer B, is there any subset $A' \subseteq A$ such that the product of the sizes of the elements in A' is exactly B?

3.1.1 A polynomial time approximation schema for knapsack problem

The only interesting case for RKM is the case where $\forall i \in I,\ a_i c_i > 0$. In fact, (i) if $a_i \geq 0$ and $c_i \leq 0$, then we can suppose that the object i does not exist, because it will never be included in any sub-optimal solution; particularly, if $a_i = 0$ and $c_i = 0$, then it is easy to see that we can delete this object without changing neither the constraint nor the value of any feasible solution i involved; (ii) if $a_i = 0$ and $c_i > 0$, or if $a_i < 0$ and $c_i \geq 0$, then every suboptimal solution contains necessarily the object i and consequently we can ignore this object and resolve the resulting subproblem. Hence, in what follows, RKM will concern the cases where $a_i c_i > 0$.

Theorem 1. *For any $\varepsilon > 0$, there is an $O(n^2)$ polynomial time approximation algorithm $\mathcal{A}$ solving every instance C of RKM with approximation ratio*

$$\frac{A(C) - \Omega(C)}{OPT(C) - \Omega(C)} \geq 1 - \varepsilon$$

where $A(C), OPT(C), \Omega(C)$ are the cardinalities of the solutions provided by $\mathcal{A}$, the optimal and the worst ones, respectively.

Proof: Throughout the proof of the theorem, we consider a fixed instance of RKM, i.e. we consider $\vec{a}$, b, $\vec{c}$ and d to be fixed; moreover, to allege the writing we will omit the fleshes over the vectors.
We define first a greedy solution for RKM.
We re-arrange the set of indices J in such a way that $\frac{a_n}{c_n} \geq \frac{a_{n-1}}{c_{n-1}} \geq \ldots \geq \frac{a_0}{c_0}$. For all integer $k \leq n$, we define the vector $\bar{x}^k$ as follows:

$$\forall j \in J,\ \bar{x}_j^0 = \begin{cases} 0 & \text{if } a_i > 0 \\ 1 & \text{if } a_i < 0 \end{cases}$$

$$\forall k > 0, \forall j \in J,\ \bar{x}_j^k = \begin{cases} 1 - \bar{x}_j^0 & \text{if } j \leq k \\ \bar{x}_j^0 & \text{if } j > k \end{cases}$$

We denote by $k(b)$ the largest index k such that $\bar{x}^k$ is feasible. The vector $\bar{x}^{k(b)}$ is called the greedy solution. Of course, $\bar{x}^0$ is the worst solution of the problem ($\Omega(I) = c \cdot \bar{x}^0$).
Without loss of generality, we can make the following hypothesis ([**H**]).
[**H**] (i) $a \cdot \bar{x}^0 \leq b$ (the set of constraints is not empty); and (ii) every $0-1$ vector deduced from $\bar{x}^0$ by changing only one of its coordinates corresponds to a feasible solution (because in the opposite case the corresponding object, being too big, is never selected and thus does not take part in any feasible solution).
We prove now that

$$c \cdot \bar{x}^{k(b)} \leq OPT(C) < c \cdot \bar{x}^{k(b)+1} \tag{1}$$

We use the following relaxation RRKM (Relaxed RKM) of RKM:

$$\text{RRKM} = \begin{cases} \max & c \cdot x & & \\ & a \cdot x & \leq & b \\ & x_i & \in & [0,1],\ \forall i \in I \end{cases}$$

In fact, we prove that the greedy solution gives an optimal solution of RRKM, or more formally, that the vector $\tilde{x}$ (called a continuous greedy solution) with

$$\begin{cases} \tilde{x}_i = \bar{x}_i^{k(b)} & \text{if } i \neq k(b)+1 \\ \tilde{x}_{k(b)+1} = \frac{b - a \cdot \bar{x}^{k(b)} + a_{k(b)+1} \bar{x}_{k(b)+1}^{k(b)}}{a_{k(b)+1}} & \end{cases}$$

where $\tilde{x}_{k(b)+1}$ is computed in such a way that $a \cdot \tilde{x} = b$ is optimal for RRKM.
In order to show that $\tilde{x}$ is optimal for the problem RRKM, we just have to prove that it is feasible and satisfies the Kuhn & Tucker conditions.

The solution $\tilde{x}$ has been computed to saturate the condition $a \cdot x \leq b$. We show now that $\tilde{x}$ is feasible, i.e. $\tilde{x}_{k(b)+1} \in [0,1]$.

We examine the following cases.
- $a_{k(b)+1} > 0$ (thus $\bar{x}^{k(b)}_{k(b)+1} = \bar{x}^{0}_{k(b)+1} = 0$).

By definition of $k(b)$, we have $a \cdot \bar{x}^{k(b)} \leq b < a \cdot \bar{x}^{k(b)+1} = a \cdot \bar{x}^{k(b)} + a_{k(b)+1}$ or $0 \leq b - a \cdot \bar{x}^{k(b)} < a_{k(b)+1}$ or $0 \leq \frac{b - a \cdot \bar{x}^{k(b)}}{a_{k(b)+1}} = x_{k(b)+1} < 1$.
- $a_{k(b)+1} < 0$ (thus $\bar{x}^{k(b)}_{k(b)+1} = \bar{x}^{0}_{k(b)+1} = 1$).

By the same arguments as previously, we have $a \cdot \bar{x}^{k(b)} \leq b < a \cdot \bar{x}^{k(b)+1} = a \cdot \bar{x}^{k(b)} - a_{k(b)+1}$ or $0 \leq b - a \cdot \bar{x}^{k(b)} < -a_{k(b)+1}$ or $0 \leq \frac{b - a \cdot \bar{x}^{k(b)} + a_{k(b)+1}}{a_{k(b)+1}} = x_{k(b)+1} < 1$.

We establish now the Kuhn & Tucker conditions. Let us consider also the problem RRKM including the box condition as a constraint denote by κ, $(\lambda_i)_{i \in I}$, $(\mu_i)_{i \in I}$, the Kuhn & Tucker coefficients.

$$\max c \cdot x$$

$$a \cdot x \leq b \tag{2}$$

$$x_i \leq 1 \tag{3}$$

$$-x_i \leq 0 \tag{4}$$

In the previous program, constraint (2) is associated to the Kuhn & Tucker coefficient κ, the set of constraints (3) to the coefficients $(\lambda_i)_{i \in I}$, and the set of constraints (4) to the coefficients $(\mu_i)_{i \in I}$.

There is a choice of κ, λ_i and μ_i which satisfies the following conditions:
(a) If $i \leq k(b)$ and $a_i > 0$ ((2) and (3) are saturated), we have $c_i = \kappa a_i + \lambda_i$. If $i \leq k(b)$ and $a_i < 0$ ((2) and (4) are saturated), we have $c_i = \kappa a_i - \mu_i$.
(b) If $i = k(b) + 1$, then $c_{k(b)+1} = \kappa a_{k(b)+1} + (\text{possibly})\lambda_{k(b)+1} - \mu_{k(b)+1}$ ((2) and possibly (3) or (4) are saturated).
(c) If $i \geq k(b) + 2$ and $a_i > 0$ ((2) and (4) are saturated), we have $c_i = \kappa a_i - \mu_i$. If $i \geq k(b) + 2$ and $a_i < 0$ ((2) and (3) are saturated), we have $c_i = \kappa a_i + \lambda_i$.

In (b), we consider $\lambda_{k(b)+1} = \mu_{k(b)+1}$. Thus $\kappa = \frac{c_{k(b)+1}}{a_{k(b)+1}} > 0$.
In (a), we determine λ_i or μ_i which are positive. The arguments: if $a_i > 0$, then $\lambda_i = c_i - \kappa a_i > 0$ as $\frac{c_i}{a_i} > \kappa$ $(a_i > 0)$; on the other hand, if $a_i < 0$, then $\mu_i = \kappa a_i - c_i > 0$ as $\frac{c_i}{a_i} > \kappa$ $(a_i < 0)$.
We use the same argument in (c).

If we note $\hat{x}$ and $\tilde{x}$ the optimal solutions of RKM and RRKM respectively, we deduce

$$c \cdot \bar{x}^{k(b)} \overset{(1)}{\leq} c \cdot \hat{x} \overset{(2)}{\leq} c \cdot \tilde{x} \overset{(3)}{\leq} c \cdot \bar{x}^{k(b)+1} \tag{5}$$

because $\bar{x}^{k(b)}$ is feasible for RKM (inequality (1) of expression (5)), $\hat{x}$ is feasible for RRKM (inequality (2) of expression (5)) and

$$c \cdot (\bar{x}^{k(b)+1} - \tilde{x}) = c_{k(b)+1}(\bar{x}^{k(b)+1}_{k(b)+1} - \tilde{x}_{k(b)+1}) \geq 0 \tag{6}$$

(inequality (3) of expression (5)).
In fact, expression (6) holds because of the following arguments: if $c_{k(b)+1} > 0$ then we have $\bar{x}^{k(b)+1}_{k(b)+1} = 1 \Longrightarrow \bar{x}^{k(b)+1}_{k(b)+1} - \tilde{x}_{k(b)+1} \geq 0$ whilst, if $c_{k(b)+1} < 0$, then we get $\bar{x}^{k(b)+1}_{k(b)+1} = 0 \Longrightarrow \bar{x}^{k(b)+1}_{k(b)+1} - \tilde{x}_{k(b)+1} \leq 0$. This concludes the proof of inequalities in expression (1).

We continue the proof of theorem 1 by proving the following inequalities.

$$0 < c \cdot \hat{x} - c \cdot \bar{x}^0 \overset{(1)}{<} c \cdot \bar{x}^{k(b)+1} - c \cdot \bar{x}^0 \overset{(2)}{\leq} 2(c \cdot \hat{x} - c \cdot \bar{x}^0) \,. \tag{7}$$

Inequality (1) in expression (7) results from expression (1).
Inequality (2) in expression (7) holds because $c \cdot \bar{x}^{k(b)+1} \leq c \cdot \bar{x}^{k(b)} + |c_{k(b)+1}| \leq c \cdot \hat{x} + |c_{k(b)+1}| \leq 2c \cdot \hat{x} - c \cdot \bar{x}^0$.
By hypothesis **[H]**, the solution $\bar{x}^0 + sign(c_{k(b)+1}) e_{k(b)+1}$ is feasible for RKM where $sign(x)$ denotes the sign of x. Thus, the fact that $c \cdot \bar{x}^0 + sign(c_{k(b)+1}).c_{k(b)+1} \leq c \cdot \hat{x}$ is equivalent to the fact $|c_{k(b)+1}| \leq c \cdot \hat{x} - c \cdot \bar{x}^0$.
This concludes the justification of inequalities in expression (7).

We show now how to obtain an $(1-\varepsilon)$-approximation.
We define RKM$^\alpha$ (see also [12,21])

$$\text{RKM}^\alpha \;=\; \left\{ \begin{array}{lll} \max & c^\alpha \cdot x & \\ & a \cdot x & \leq \; b \\ & x_i & \in \; \{0,1\} \end{array} \right.$$

with $c_i^\alpha = \left\lfloor \frac{c_i}{\alpha} \right\rfloor$ and $\alpha = \frac{c \cdot \bar{x}^{k(b)+1} - c \cdot \bar{x}^0}{2n} \varepsilon > 0$, where for every $x \in I\!R$, $\lfloor x \rfloor$ denotes the floor of x.
Since solutions $\hat{x}$ and $\hat{x}^\alpha$ are feasible for RKM and RKM$^\alpha$ respectively, we have $c^\alpha \cdot \hat{x} \leq c^\alpha \cdot \hat{x}^\alpha$ and $c \cdot \hat{x}^\alpha \leq c \cdot \hat{x}$.
On the other hand, we have also $0 \leq c_i - \alpha c_i^\alpha < \alpha$ and thus according to expression (1),

$$\alpha c^\alpha \cdot \hat{x} \leq \alpha c^\alpha \cdot \hat{x}^\alpha \leq c \cdot \hat{x}^\alpha \leq c \cdot \hat{x} \leq c \cdot \bar{x}^{k(b)+1} \tag{8}$$

and

$$0 \leq c \cdot \hat{x} - c \cdot \hat{x}^\alpha \leq c \cdot \hat{x} - \alpha c^\alpha \cdot \hat{x} = (c - \alpha c^\alpha) \cdot \hat{x} \leq n\alpha. \tag{9}$$

We have also by expressions (7) and (8): $\frac{c \cdot \hat{x} - c \cdot \hat{x}^\alpha}{c \cdot \hat{x} - c \cdot \bar{x}^0} \leq \frac{2n\alpha}{c \cdot \bar{x}^{k(b)+1} - c \cdot \bar{x}^0} = \varepsilon$ or $\frac{c \cdot \hat{x}^\alpha - c \cdot \bar{x}^0}{c \cdot \hat{x} - c \cdot \bar{x}^0} \geq 1 - \varepsilon$, and thus we can deduce so that $\hat{x}^\alpha$ is an ε-approximation for RKM.

It remains to show that there is a polynomial time algorithm which determines an exact solution of RKM$^\alpha$.
We consider RKm$^\alpha_d$ defined as follows:

$$\text{RKm}^\alpha_d \;=\; \left\{ \begin{array}{lll} \min & a \cdot x & \\ & c^\alpha \cdot x & \geq \; d \\ & x_i & \in \; \{0,1\} \end{array} \right.$$

We have $c^\alpha \cdot \hat{x}^\alpha - c^\alpha \cdot \bar{x}^0 \leq \frac{c \cdot \hat{x}^\alpha}{\alpha} - c^\alpha \cdot \bar{x}^0 < \frac{c \cdot \bar{x}^{k(b)+1}}{\alpha} - c^\alpha \cdot \bar{x}^0 < \frac{2n}{\varepsilon} + \frac{c \cdot \bar{x}^0}{\alpha} - c^\alpha \cdot \bar{x}^0 \leq \frac{2n}{\varepsilon} + n$, the first inequality holding because of (8) and (9), and the second one holding because of

expression (1).
If we put $\bar{d} = \lfloor n(1+\frac{2}{\varepsilon})\rfloor$, we have also $0 \le c^\alpha \cdot \hat{x}^\alpha - c^\alpha \cdot \bar{x}^0 \le \bar{d}$.

We continue the proof of theorem 1 by proving the following emphasized proposition:
Let $\mathcal{D} = \{c^\alpha \cdot \bar{x}^0, c^\alpha \cdot \bar{x}^0 + 1, \ldots, c^\alpha \cdot \bar{x}^0 + \bar{d}\} \subset I\!N$. Once an optimal solution of RKm_d^α is known for every $d \in \mathcal{D}$, we can determine an optimal solution of RKM^α which constitutes an $(1-\varepsilon)$-approximate solution for RKM.
First, given an instance of a problem, let us use the expression "value of an instance" to denote "the value of the optimal solution of the instance". Also, given an instance of RKm_d^α, let us denote by $v(d)$ its value. Finally, let d_{max} the maximum of the set $\mathcal{D}_{max} = \{d \in \mathcal{D} : v(d) \le b\}$. We note that $\mathcal{D}_{max} \neq \emptyset$ because $c^\alpha \cdot \bar{x}^0 \in \mathcal{D}_{max}$ ($a \cdot \bar{x}^0 < b$, by hypothesis [H]).
If $\check{x}$ is an optimal solution of the instance of RKm_d^α where $d = d_{max}$, then $\check{x}$ is the searched solution of RKM^α. The arguments: obviously, $a \cdot \check{x} \le b$; if $d_{max} = c^\alpha \cdot \bar{x}^0 + \bar{d}$, then as $a \cdot \check{x} \le b$, by definition of $\hat{x}^\alpha$ we have $c^\alpha \cdot \hat{x}^\alpha \ge c^\alpha \cdot \check{x}$; but $d_{max} = c^\alpha \cdot \bar{x}^0 + \bar{d} \le c^\alpha \cdot \check{x} \le c^\alpha \cdot \hat{x}^\alpha \le c^\alpha \cdot \bar{x}^0 + \bar{d}$ thus $c^\alpha \cdot \check{x} = c^\alpha \cdot \hat{x}^\alpha$, i.e. $\check{x}$ is optimal for RKM^α; on the other hand, if $d_{max} \neq \bar{c}^\alpha \cdot x^0 + \bar{d}$, then let x be such that $c^\alpha \cdot x > c^\alpha \cdot \check{x} \ge d_{max} + 1$; this implying that $a \cdot x$ is greater than the value of the instance of RKm_d^α with $d = d_{max} + 1$ (x is feasible), thus $a \cdot x > b$ by definition of d_{max}.
¿From the arguments presented above, we can deduce that $\check{x}$ is an optimal solution of RKM^α and conclude the proof of the emphasized proposition.

To conclude the proof of the theorem, we have to show that it is possible to find an optimal solution z of RKm_d^α for every instance such that $d \in \mathcal{D}$ in polynomial time.
In the case of real coefficients, we cannot deal directly with dynamic programming arguments. So we have to proceed in two stages.

We start by solving the cases where all the coefficients are positive; we consider then the following problem

$$|\mathrm{RKm}^\alpha|_\delta^k = \left\{ \begin{array}{lll} \min & \sum_{i=1}^{k} |a_i| x_i & \\ & \sum_{i=1}^{k} |c^\alpha| x_i & \ge \delta \\ & x_i & \in \{0,1\} \end{array} \right.$$

We show that we can resolve problem $|\mathrm{RKm}^\alpha|_\delta^n$ in $n(\bar{d}+1)$ steps, for $\delta \in \{0, \ldots, \bar{d}\}$, i.e. solve $|\mathrm{RKm}^\alpha|_{d-c^\alpha \cdot \bar{x}^0}^n$ for $d \in \mathcal{D}$.
We note here that, in order to determine the solutions of $|\mathrm{RKm}^\alpha|_\delta^{k+1}$ for $\delta \in \{0, \ldots, \bar{d}\}$, we only have to know the solutions of $|\mathrm{RKm}^\alpha|_\delta^k$ for all $k \le n$ and $\delta \in \{0, \ldots, \bar{d}\}$. In fact, for all $\delta \in \{0, \ldots, \bar{d}\}$, to determine the better solution of $|\mathrm{RKm}^\alpha|_\delta^{k+1}$ with $x_{k+1} = 0$, it is sufficient to solve problem $|\mathrm{RKm}^\alpha|_\delta^k$. Also, to determine the better solution of $|\mathrm{RKm}^\alpha|_\delta^{k+1}$ with $x_{k+1} = 1$, we only have to solve problem $|\mathrm{RKm}^\alpha|_{(\delta - |c_{k+1}^\alpha|)^+}^k$, where given a number x, $x^+ = \max\{x, 0\}$. If we denote by v_1 the value of the former and by v_2 the value of the latter, then by the dynamic programming principle, the relative position of the numbers v_1 and $v_2 + |a_{k+1}|$ determines the searched solution. Thus, by induction, if we can resolve $|\mathrm{RKm}^\alpha|_\delta^n$ for $n = 1$, we also can resolve it for all n.

As $\bar{d} \leq n(1+\frac{2}{\varepsilon})$, the number of necessary steps is $O(n^2)$.

We prove now that the first stage is sufficient to find a solution of RKm_d^α for every instance in which $d \in \mathcal{D} = \{c^\alpha \cdot \bar{x}^0, \ldots, c^\alpha \cdot \bar{x}^0 + \bar{d}\}$.
Consider a solution y of $|\mathrm{RKm}^\alpha|^n_{d-c^\alpha \cdot \bar{x}^0}$. We set: $z_i = y_i + \bar{x}_i^0$, if $a_i > 0$, $z_i = -y_i + \bar{x}_i^0$, if $a_i < 0$. We have also $y = |z - \bar{x}^0|$. It is easy to see that z is an optimal solution of RKm_d^α. Indeed, $\forall x \in \{0,1\}^n$ we have $|x - \bar{x}^0| \in \{0,1\}^n$. Thus z is an optimal solution of

$$\begin{cases} \min & |a| \cdot |x - \bar{x}^0| & & \\ & |c^\alpha| \cdot |x - \bar{x}^0| & \geq & d - c^\alpha \cdot \bar{x}^0 \\ & x_i & \in & \{0,1\} \end{cases}$$

Moreover, $\forall i \in I, \forall x \in \{0,1\}^n$ and by definition of $\bar{x}_i^0$ we have, $a_i(x_i - \bar{x}_i^0) = |a_i||(x_i - \bar{x}_i^0)|$ and $c_i^\alpha(x_i - \bar{x}_i^0) = |c_i^\alpha||(x_i - \bar{x}_i^0)|$. Thus z is also an optimal solution of the problem

$$\begin{cases} \min & a \cdot (x - \bar{x}^0) & & \\ & c^\alpha \cdot (x - \bar{x}^0) & \geq & d - c^\alpha \cdot \bar{x}^0 \\ & x_i & \in & \{0,1\} \end{cases}$$

which has the same solutions as the problem

$$\mathrm{RKm}_d^\alpha \quad = \quad \begin{cases} \min & a \cdot x & & \\ & c^\alpha \cdot x & \geq & d \\ & x_i & \in & \{0,1\} \end{cases}$$

And this concludes the proof of theorem 1. ∎

The above proof shows, quite clearly, how the ratio has to be changed to preserve the result. In particular, the expression $A(I) - \Omega(I)$ seems to be very convenient for the proof of the expression (7) and also in the definition of the greedy solution of RKM. This expression is crucial for the transformation from RKM to KM.

3.1.2 Some consequences of knapsack theorem

Theorem 1 allows us to bring to the fore a pseudo-polynomial algorithm for the following generalization of KM.

GKM. Given two positive integers b and b', given a real vector $\vec{a}$, and a vector $\vec{c}$ with elements c_i such that $c_i \in I\!N^* \cup -I\!N^*$, $\sum_{c_i>0} c_i \leq b$ and $\sum_{c_i<0} c_i \geq -b'$, solve the corresponding RKM.

Corollary 1. *There is an exact $O(2(b+b')n^2))$ algorithm for GKM.*

In fact, if we select $\varepsilon = \frac{1}{2(b+b')}$, the schema constructed in theorem 1 works in $O(2(b+b')n^2))$ and gives an optimal solution.
Indeed, we have $\Omega(I) \geq -b'$ and so $\frac{A(I)-\Omega(I)}{OPT(I)-\Omega(I)} \geq 1 - \varepsilon$ or $0 \leq OPT(I) - A(I) \leq \varepsilon(OPT(I) - \Omega(I)) \leq \varepsilon(b+b') \leq \frac{1}{2}$. Thus, as $OPT(I)$ and $A(I)$ are integers, they are

equal.
So, we can see how the hypothesis *c has integer coordinates* intervenes and why RKM has not a similar pseudo-polynomial algorithm.

We consider now the following generalization GS of S.

GS. Given two positive integers M and N, given a vector $(c_i)_{i\in\{1,\dots,n\}}$ with $c_i \in I\!N^* \cup -I\!N^*$, and $\sum_{c_i<0} c_i \geq -N$, does a (0,1) vector x exist such that $c \cdot x = M$?

Corollary 2. *GS admits an $O(2(M+N)n^2)$ algorithm.*

By solving the problem:

$$\begin{cases} \max & \vec{c}\cdot\vec{x} & & \\ & \vec{c}\cdot\vec{x} & \leq & M \\ & x_i & \in & \{0,1\},\ \forall i \in I \end{cases}$$

we can answer the question if x exists.
But we note that, to solve in such a way the subset product problem ([10]), we have to consider the following instances of RKM:

$$\begin{cases} \max & \ln(\vec{c})\cdot\vec{x} & & \\ & \ln(\vec{c})\cdot\vec{x} & \leq & \ln(M) \\ & x_i & \in & \{0,1\},\ \forall i \in I \end{cases}$$

Those instances are not included in the particular case of RKM for which we give a pseudo-polynomial algorithm.
This illustrates why SP is NP-complete in the strong sense ([10]).

3.1.3 Some equivalence results

Proposition 2. *RKM and RKm are equivalent.*

Proof: If we denote by $\bar{x}^0$ the worst solution, we construct from every instance

$$\text{IM} = \begin{cases} \max & \vec{c}\cdot\vec{x} & & \\ & \vec{a}\cdot\vec{x} & \leq & b \\ & x_i & \in & \{0,1\},\ \forall i \in I \end{cases}$$

of RKM, the following instance

$$\text{Im}\ (=\propto_{\text{RKM,RKm}}(\text{IM})) = \begin{cases} \min & \vec{c}\cdot\vec{y} & & \\ & \vec{a}\cdot\vec{y} & \geq & \vec{a}\cdot\vec{1} - b \\ & y_i & \in & \{0,1\},\ \forall i \in I \end{cases}$$

of RKm. The algorithm $\psi_{\text{RKm,RKM}}$ (section 2) can be defined as $\psi_{\text{RKm,RKM}}(\vec{y}) = \vec{1} - \vec{y}$ for all instances IM and Im, for all feasible vectors $\vec{y}$, of Im.

As soon as $\vec{a} \cdot \vec{y} \geq \vec{a} \cdot \vec{1} - b$, we have obviously $\vec{a} \cdot (\vec{1} - \vec{y}) \leq b$, i.e. $\psi_{\text{RKm,RKM}}(\vec{x})$ is feasible for IM.
The linear function $f_{\text{Im,IM}}$ is defined by $f_{\text{Im,IM}}(\vec{z}) = \vec{c} \cdot \vec{1} - \vec{z}$.
Also, from every instance of RKm

$$\text{Im} = \begin{cases} \min & \vec{c} \cdot \vec{y} \\ & \vec{a} \cdot \vec{y} \geq d \\ & y_i \in \{0,1\},\ \forall i \in I \end{cases}$$

we construct the following instance of RKM

$$\text{IM} = \begin{cases} \max & \vec{c} \cdot \vec{x} \\ & \vec{a} \cdot \vec{x} \leq \vec{c} \cdot \vec{1} - b \\ & x_i \in \{0,1\},\ \forall i \in I \end{cases}$$

in such a way that $\propto_{\text{RKm,RKM}} \circ \propto_{\text{RKM,RKm}} = Id$ and $\propto_{\text{RKM,RKm}} \circ \propto_{\text{RKm,RKM}} = Id$, where Id is the identity function.
We define $f_{\text{IM,Im}} = f^{-1}_{\text{Im,IM}}$ and $\psi_{\text{RKM,RKm}}(\vec{x}) = \vec{1} - \vec{x}$ which satisfies the property (i) of definition 3. Moreover, $\forall$Im, $\forall$IM, $(\psi_{\text{RKm,RKM}} \circ \psi_{\text{RKM,RKm}})|_{C_{\text{Im}}} = Id_{C_{\text{Im}}}$ and $(\psi_{\text{RKM,RKm}} \circ \psi_{\text{RKm,RKM}})|_{C_{\text{IM}}} = Id_{C_{\text{IM}}}$. ∎

Proposition 3. *KM and Km are equivalent, provided that we impose for each problem an additional condition: we just consider the instances of KM where $\vec{a} \cdot \vec{1} \leq b$ and the ones of Km where $\vec{a} \cdot \vec{1} \geq d$.*

We note that this condition restrict us to the interesting cases. Indeed, this condition means that the set of feasible vectors of Im is not empty and that $\vec{1}$ is not feasible for KM which, in the opposite case, would be a trivial solution.
With this restriction, the reduction described in proposition 2 is valid again.

Proposition 4. *The two following problems are weakly equivalent:*
(i) RKM with the additional condition that the set of feasible solutions is not empty (that means, with the notations defined above, that the worst solution $\vec{x^0}$ is feasible).
(ii) PRKM (positive real knapsack maximization problem), which is the previous problem with positive coefficients for the objective function, the constraint function and the second member.

Proof: We follow here a reasoning similar to the one of the last step of theorem 1. Let us consider an instance of the former:

$$\text{IM} = \begin{cases} \max & \vec{c} \cdot \vec{x} \\ & \vec{a} \cdot \vec{x} \leq b \\ & x_i \in \{0,1\},\ \forall i \in I \end{cases}$$

The feasibility of $\vec{x^0}$ is written $\vec{a} \cdot \vec{x^0} \leq b$.
We construct the following instance of the latter:

$$|\text{IM}| = \begin{cases} \max & |\vec{c}| \cdot \vec{x} \\ & |\vec{a}| \cdot \vec{x} \leq b - \vec{a} \cdot \vec{x^0} \\ & x_i \in \{0,1\},\ \forall i \in I \end{cases}$$

This construction defines a reduction $\propto_{\text{RKM,PRKM}}$.
The function $\psi_{\text{PRKM,RKM}}$ is defined as follows.
For all IM and $|\text{IM}| = \propto_{\text{RKM,PRKM}}(\text{IM})$ and for all $x \in C_{|\text{IM}|}$

$$(\psi_{\text{PRKM,RKM}}(x))_i = \begin{cases} x_i + \bar{x}_i^0 & \text{if} \quad a_i > 0 \\ -x_i + \bar{x}_i^0 & \text{if} \quad a_i < 0 \end{cases}$$

i.e., $\forall \vec{x} \in C_{|\text{IM}|}$, $|\psi_{\text{PRKM,RKM}}(\vec{x}) - \vec{\bar{x}^0}| = x$. In fact, by definition of $\vec{\bar{x}^0}$, we have $\vec{a} \cdot (\psi_{\text{PRKM,RKM}}(\vec{x}) - \vec{\bar{x}^0}) = |\vec{a}| \cdot |\psi_{\text{PRKM,RKM}}(\vec{x}) - \vec{\bar{x}^0}|$ and consequently $\psi_{\text{PRKM,RKM}}(\vec{x}) \in C_{\text{IM}}$.
Let us define $f_{|\text{IM}|,\text{IM}}(\vec{z}) = \vec{z} - \vec{ax} \cdot \vec{\bar{x}^0}$. By the same argument, $\vec{x} \cdot (\psi_{\text{PRKM,RKM}}(\vec{x}) - \vec{\bar{x}^0}) = |\vec{c}| \cdot |\psi_{\text{PRKM,RKM}}(\vec{x}) - \vec{\bar{x}^0}|$ thus, $f_{|\text{IM}|,\text{IM}}$ satisfies the conditions of definition 3 of the weak equivalence.
Since the second problem can be considered as a sub-problem of the former, every application $\propto_{\text{PRKM,RKM}}$, as well as every application $\psi_{\text{PRKM,RKM}}$, are the identical applications on the convenient spaces, and moreover, for all instance PIM of PRKM, $f_{\text{PIM,PIM}} = Id_{\mathbb{R}}$.
∎

3.2 Approximation results for set covering, hitting set, vertex covering, independent set, clique, set packing

Let us take now VC and IS. In fact, there is a bijective affine transformation of the objective function of VC to the objective function of IS if we take $\lambda = -1$ and $\mu = n$, where n is the order of the graph; hence, the equivalence defined in section 2 together with the approximation ratio defined there, make vertex covering and independent set to be equivalent with respect to approximation ratio of an algorithm supposed to solve both problems. On the other hand, the way we have defined the new approximation ratio implies that this ratio is similar to the old one for the case of IS.
Recently, some researchers ([2]), have proved the following:

Proposition 5. ([2]). *VC, even for bounded degree graphs and IS for bounded degree graphs*[2]*, do not admit a polynomial time approximation* **schema** *unless $P = NP$.*

Based on this result and on our definition of equivalence, we can conclude the following more strict result for the approximation measure proposed here:

Theorem 2. *VC and IS, even for bounded degree graphs*[2]*, do not admit a polynomial time approximation* **algorithm** *with a constant approximation ratio (independent of the bound) unless $P = NP$.*

Let us now consider some other combinatorial problems like maximum set packing, maximum clique and maximum complete bipartite subgraph.

[2]It is well known that IS for bounded degree graphs admit polynomial time approximation algorithms with ratios depended on the bounds

Given a collection of finite sets $\mathcal{S}$ ($|\mathcal{S}| = n$), a packing is a subcollection $\mathcal{S}' \subseteq \mathcal{S}$, all of whose members are mutually disjoint and the maximum set packing problem is to find a packing of maximum size.
Given a graph $G = (V, E)$, a clique of G is a subset $V' \subseteq V$ such that every two vertices of V' are joined by an edge in E and the maximum clique problem is to find a maximum size set V' inducing a clique in G (a maximum size clique); the maximum complete bipartite subgraph problem consists in finding a set $U \subseteq V$ of maximum size inducing a complete bipartite subgraph.

Given two problems Π, Π', we call them approximate equivalent if any approximation algorithm solving the former solves also the latter and vice versa within the same approximation ratio.
Well known results ([10,15,17,22]) prove the approximate equivalence between IS, set packing, clique and complete bipartite subgraph; this equivalence is also true in the framework of our approach, since all these problems are maximization problems with $\Omega(I) = 0$.

Let us now consider the case of SC. Given a collection $\mathcal{S}$ ($|\mathcal{S}| = n$) of subsets of a finite set C ($|C| = m$), a cover is a subcollection $\mathcal{S}' \subseteq \mathcal{S}$ such that $\bigcup_{s_i \in \mathcal{S}'} s_i = C$ and the minimum set cover problem (SC) is to find a cover of minimum size. Every instance I of SC can be represented by a bipartite graph $B = (S, C, E)$, called the characteristic graph B of I, where the vertex set S denotes the family $\mathcal{S}$ and the vertex set C the elements of the set C. The graph B contains the edge $s_i c_j$ if $c_j \in S_i$.
Every instance I of VC is expressed in terms of a graph $G = (V, E)$, which can equivalently be represented by a bipartite graph $B_G = (V, E, E')$ where, E', which is the edge set of B_G, contains the pairs $v_i e_j$ such that e_j is incident to v_i in G.
Clearly, the instances of VC are exactly the instances of SC where every element of C is contained in exactly two subsets of $\mathcal{S}$. Hence, we can treat every instance of VC as an instance of the SC, by considering G or equivalently B_G as the characteristic graph of this instance. Consequently, as VC is a subproblem of SC, the negative result of theorem 2 holds also for SC.
Let us now recall the definition of another NP-complete problem, the minimum hitting set problem. Let n and m denote two integers and $\mathcal{S} = \{S_1, ..., S_n\}$ a family of n subsets drawn from an m-element set C; a subset $H \subseteq C$ is a hitting set for the family $\mathcal{S}$ if H has a non-empty intersection with each element of this family; the minimum hitting set problem is to find a hitting set of minimum cardinality. It is easy to see that hitting set is approximate equivalent to SC, because a SC becomes hitting set via the interchanging of the two sets $\mathcal{S}, C$.

All the above discussion is resumed, for the case of our approach, in the following theorem:

Theorem 3. *The problems SC, VC, IS, clique, set packing, complete bipartite subgraph, do not have a polynomial time approximation algorithm with a ratio independent of the size of the instance unless $P = NP$.*

For SC however, we present a constant ratio polynomial time approximation algorithm for some restricted instances. This algorithm works on the C-intersection graph $G_C = (C, E)$

```
begin
    S' ← ∅;
    construct G_C;
    obtain a maximum matching M in G_C;
    S' ← S' ∪ {set of labels on the edges of M};
    C ← C \ {c_k, c_l : c_k c_l ∈ M}
    while C ≠ ∅ do
        choose a vertex c_j ∈ C;
        choose an edge e incident to c_j;
        C ← C \ {c_j};
        S' ← S' ∪ {set of labels of e}
    od
    remove the (eventual) duplications from S'
end.
```

Algorithm 1.

of SC defined as follows.

Definition 4. The C-intersection graph $G_C = (C, E)$ of SC is an edge-labelled graph whose vertices are the elements of C, two points c_i, c_j are adjacent iff $\exists k : \{c_i, c_j\} \subseteq S_k$, and the label of such edge $c_i c_j$ is s_k. ∎

Theorem 4. *Algorithm 1 is an $O(m^{2.5})$ heuristic which, if $m < n$, approximates SC with ratio $\frac{1}{2}$.*

Proof: Every vertex j of G_C represents the element j of the set C, and labels of the edges incident to j represent exactly all the subsets containing the element j. Thus, each time algorithm adds the labels of an edge e in S', in fact it covers the elements represented by the endpoints of e by the subsets whose names are labels of e. During this phase, m' edges are selected, where $m' = |M|$. This entails the removal of $2m'$ vertices of C.
Now in the set C, remain $m - 2m'$ vertices called exposed vertices. This set is an independent set because, if there were an edge e' between two vertices, then e' would be added in M. For each such vertex c_j, we add in S' the labels of one of the edges incident to c_j. Thus, after removing duplications, the set S' has cardinality at most $m' + (m - 2m') = m - m'$. Let us denote by σ the cardinality of S'.
The fact that the exposed vertices of C associated to M form an independent set, means in terms of SC that there is no subset covering two elements represented by two vertices of this set. Thus, the optimum solution of the problem contains at least as many subsets as the cardinality of the set of the exposed vertices which is equal to $m - 2m'$. Let us denote by β the cardinality of the optimal solution for SC.
On the other hand, we take $\Omega(I) = |\mathcal{C}| = m$ (we discuss later this choice).

We have then: $\rho = \frac{\Omega(I)-\sigma}{\Omega(I)-\beta} \geq \frac{m-m+m'}{m-m+2m'} = \frac{m}{2m}$, thus $\rho \geq \frac{1}{2}$.
It is easy to see that the more "expensive" operation of the algorithm is the maximum matching ([9,15]) performed in $O(c^{2.5})$ and this determines its complexity, and this completes the proof of the theorem. ∎

The example of SC is an interesting example of how the choice of the worst value intervenes in the evaluation of the approximation performance of an algorithm. In fact, SC is a problem of choosing a number of sets having a certain property and as the union of the elements of $\mathcal{S}$ equals C, the whole $\mathcal{S}$ (with cardinality n) appears to be a natural worst value solution. On the other hand, it is equally natural to suppose that one does never have a solution where the sets chosen are more than the elements to cover, hence, another feasible solution for the problem is the one obtained by an arbitrary choice of a set per element, and the value of this solution is m (the cardinality of C); the so obtained solution can equally be considered as the worst value solution. It is easy to see that if we systematically choose m as the value of the worst solution, then algorithm 1 is always $\frac{1}{2}$-approximation. But, trying to be coherent with our principal motivation, that is, adopting an approximation measure rich in information about problem's approximability, we have chosen $\min\{n, m\}$ as worst value. Obviously this choice makes SC more difficult than the a priori choice of m to be always the worst value because if $\min\{n, m\} = n$, then algorithm 1 has no constant approximation ratio.

4 Discussion and final remarks

We have tried to point out some strange side effects of the usual approximation theory, due to the fact that it is conceived without taking into account requirements of the optimization theory. We have also tried to capture the complexity and the lack of a concise definition of a consistent polynomial approximation theory, respecting all the alternative (equivalent) ways an optimisation problem can be expressed as an integer program and more particularly the affine transformation of the objective function. We have firstly defined an equivalence among optimization problems, context strongly necessary for the definition of an approximation theory. After that, we have searched for a measure which respects the equivalence of pairs of problems under affine transformation of their objective functions. We have imposed as principal axiom on the approximation the respect of this equivalence and we have proved that the approximation ratio as a two-variable function cannot verify this axiom. We have then defined a three-variable function as a new approximation ratio which is coherent to the defined equivalence and under the choice of the variables the new ratio is introduced by an axiomatic approach. The consideration of solutions other than the optimal ones seems to be the least thing an approximation theory has to take into account, given that when we study the deterministic approximation performance of an algorithm, we always deal with the worst case. Moreover, even for practical reasons, when dealing with "real-life" applications, the quality of an approximation algorithm is measured on both its approximation performance and on the quality of the solutions provided by it. For this last point, the taking away from the worst feasible solution is a reasonable prerequisite for an algorithm.

In fact, the requirement of a coherent approximation ratio is very strong. Many researchers try to define subclasses of NP-complete problems having a kind of continuity property in the sense that the design of a polynomial time approximation algorithm (schema) for one of the problems contained in such a class imply the design of polynomial time approximation algorithm (schema) with asymptotically the same approximation ratio for every other problems of the class ([16,19,22]). We think that a prerequisite for the efficient definition of such classes intersects the concise definition of a consistent approximation theory respecting the continuity in the interior of a given problem.

Indeed, the concept of the affine transformation is not only mathematically challenging, but also interesting from a "computer science" point of view. Think that two (perhaps the most) famous NP-complete problems, the vertex covering and the independent set, are equivalent in the sense of the equivalence we have defined in definition 2. But the existence of a 2-approximation algorithm (following the classical approach) for the former does not imply an approximation algorithm with a bounded error for the latter. Of course, a generalization of this work could be the conception of a measure valid for a large class of transformations. But, the concept of affine transformations is already very natural and resides in many real applications where tools of the operational research and the combinatorial optimization have to be used.

We think that a large part of the inconsistencies of the classical approximation theory is hidden behind the conception of the notion of the approximation algorithm. When someone deals with a maximization problem, the notion of going away from the worst case solution Ω is included there by default. On the contrary, in a minimization problem, an eventual approximation algorithm as the ones known up to now, does not incorporate this prerequisite in its conception. The result of such a lack is evident in the case of vertex covering where, as we have already mentioned, a solution including all the vertices of the graph may be (and in fact it is in the most of instances) a good solution.
It is obvious that the removal of the inconveniences and the success of the tentative on designing a consistent approximation theory do inevitably treat the changing of the conception of an approximation algorithm. Moreover, this new conception has to take into account parameters other than the value of the sub-optimal solution found by the algorithm (why not value of the worst case solution?).

We have to note here that the notion of equivalence we have introduced and used throughout this paper, reminds slightly the notion of L-reduction introduced by Papadimitriou and Yannakakis in [16] whose goal is to preserve the polynomial time approximation schemata between problems linked by such a reduction. Our notion of equivalence is however quite different. In fact, among the properties of our reduction is the preservation of any value of approximation ratio, not only the particular value $1 \pm \epsilon$, for an ϵ arbitrarily small.
On the other hand, the L-reduction, conceived for the usual approximation ratio, includes some of the inconveniences mentioned, as (i) the dissymmetry between maximization and minimization problems, and (ii) bad behaviour over affine transformation.

In fact, let us consider the two following problems:

$$\Pi = \begin{cases} \max\ c \cdot x \\ x \in C \end{cases} \qquad \Pi' = \begin{cases} \max\ \lambda c \cdot x + \mu \\ x \in C \end{cases}$$

As soon as we allow instances with $OPT(I)$ arbitrarily close to 0, there is an L-reduction from Π to Π' only if μ is negative. Therefore, if we consider the equivalence relation $\mathcal{R}_L$ defined on two problems Π and Π' (via the L-reduction) by $\Pi \mathcal{R}_L \Pi' \iff (\Pi\ L -$ reduces to $\Pi') \wedge (\Pi' L -$ reduces to $\Pi)$, then Π and Π' are equivalent only if $\mu = 0$. This fact signifies that the notion of affine transformation is incompatible with L-reduction. However, affine transformation is very natural particularly in the case of the connection between vertex covering and independent set.

We have wished to point out some inconsistencies due to the lack of a formal (axiomatized) approximation theory and we have tried to propose definitions for some notions "sine qua non" for such a theory. We hope that this issue will receive comments and answers and a fruitful discussion on this matter will be open.

Acknowledgements. We wish to thank our colleagues, members of the *Groupe de Recherche en Optimisation Combinatoire et en Informatique Théorique* of LAMSADE, University of Paris-Dauphine, for their helpful comments and discussions. More particularly, we wish to thank Christian Michelot for our fruitful discussions and for having put in our disposal his intuition and his knowledge on linear programming, and Pascal Grisoni for his help to clarify some notions related to the concept of the worst case solution. We also thank Michel Minoux for helpful comments and discussions on this work.

References

[1] A. Aiello, E. Burattini, A. Massarotti and F. Ventriglia, *A New Evaluation Function for Approximation Algorithm*, Sem. IRIA, 1977.

[2] S. Arora, C. Lund, R. Motwani, M. Sudan and M. Szegedy, *Proof Verification and Intractability of Approximation Problems*, Proc. of 33rd IEEE Symposium on Foundations of Computer Science FOCS'92, pp. 14-23, 1992.

[3] G. Ausiello, A. D'Atri and M. Protasi, *Structure Preserving Reductions among Convex Optimization Problems*, J. Comput. System Sci. vol 21, pp. 136-153, 1980.

[4] C. Berge, *Graphs and Hypergraphs*, North Holland, Amsterdam, 1973.

[5] B. Bollobás, *Random Graphs*, Academic Press, London, 1985.

[6] M. Demange et V. Th. Paschos, *Quelques Etapes vers la Conciliation de la Théorie d'Approximation et celle d'Optimisation: Une Nouvelle Théorie d'Approximation Polynomiale et Résultats Préliminaires*, note to appear in the "Comptes Rendus de l'Académie des Sciences", Paris, 1993.

[7] M. Demange, P. Grisoni and V. Th. Paschos, *Quelques Résultats dans le Cadre d'une Nouvelle Théorie de l'Approximation Polynomiale*, in French, 1993.

[8] M. Demange, *Optimisation, Approximation, Complexité: Une Nouvelle Approche*, PhD Dissertation, in preparation.

[9] J. Edmonds and R. M. Karp, *Theoretical Improvements in Algorithmic Efficiency for Network Flow Problems*, J. ACM vol. 19(2), pp. 248-264, 1972.

[10] M. R. Garey and D. S. Johnson, *Computers and Intractability. A Guide to the Theory of NP-Completeness*, W. H. Freeman and Company, San Francisco, 1979.

[11] M. Gondran et M. Minoux, *Graphes et Algorithmes*, eds. Eyrolles, Paris, 1985.

[12] O. H. Ibarra and C. E. Kim, *Fast Approximation Algorithms for the Knapsack and Sum of Subset Problems*, J. of ACM vol. 22(4), pp. 463-468, 1975.

[13] H. R. Lewis and C. H. Papadimitriou, *Elements of the Theory of Computation*, Prentice Hall, New Jersey, 1981.

[14] G. L. Nemhauser, L. A. Wolsey and M. L. Fischer, *An Analysis of Approximations for Maximizing Submodular Set Functions*, Math. Program. vol 14, 1978.

[15] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*, Prentice Hall, New Jersey, 1981.

[16] C. H. Papadimitriou and M. Yannakakis, *Optimization, Approximation and Complexity Classes*, J. Comput. System Sci. vol. 43, pp. 425-440, 1991.

[17] V. Th. Paschos, *A "Tour d'Horizon" about how Optimal Solutions to Set Cover, Set Packing, Hitting Set, Vertex Cover, Independent Set and Clique can be Approximated*, manuscript, 1992.

[18] V. Th. Paschos and L. Renotte, *Approximability Preserving Reductions for NP-Complete Problems*, cahier du LAMSADE #112, 1993.

[19] L. Renotte, *Sur les Réductions Préservant la Sous-Optimalité*, Mémoire de DEA, Université de Paris I - Sorbonne, 1991.

[20] S. Sahni and T. Gonzales, P-Complete Approximation Problems, J. ACM vol. 23(3), 1976.

[21] M. Sakarovitch, *Optimisation Combinatoire, Programmation Discrète*, Hermann, Paris, 1984.

[22] H. U. Simon, *On Approximate Solutions for Combinatorial Optimization Problems*, SIAM J. of Disc. Math. vol. 3(2), pp. 294-310, 1990.