(*AI EDAM*) (1991) **5**(3), 137–171

AUTOMOBILE TRANSMISSION DESIGN AS A CONSTRAINT SATISFACTION PROBLEM: MODELLING THE KINEMATIC LEVEL

Bernard A. Nadel and Jiang Lin

Computer Science Department, Wayne State University, Detroit, MI 48202, U.S.A.

This paper describes our preliminary results in applying constraint satisfaction techniques in a system we call TRANS-FORM for designing automatic automobile power transmissions. The work is being conducted in collaboration with the Ford Motor Company Advanced Transmission Design Department in Livonia, Michigan. Our current focus is on the design of the mechanical subsystem, but we anticipate extending this later to the electrical and hydraulic subsystems also. For simplicity, in the initial work reported here we restrict ourselves to the relatively well-explored class of transmissions having four forward speeds and one reverse speed, built from two planetary gearsets, cross-connected by two permanent links. Moreover, we pursue design of such transmissions only at the 'kinematic level'. These two restrictions correspond to limiting respectively the *breadth* (generality) and the *depth* (detail or granularity) of the search space employed. We find that, at least for the restricted version of the problem pursued here, transmission design is an application very naturally formulated as a constraint satisfaction problem. Our present problem requires only 10 variables, with an average of about seven values each, and 43 constraints—making it similar in difficulty to about the 10-queens problem. So far, two of the classic transmissions, known as Axod and HydraMatic, have been rediscovered (at the kinematic level) by our program. Preliminary results also indicate that the constraint satisfaction framework will continue to remain adequate and natural even when the search space is allowed to be much broader and deeper. We expect that searches of such expanded spaces will soon lead to the discovery of totally new transmissions.

1. Introduction

The Constraint Satisfaction Problem (CSP) is ubiquitous in Artificial Intelligence. It has received intense study from many researchers, as seen for example in Fikes (1970), Waltz (1975), Gaschnig (1974, 1977, 1978, 1979), Rosenfeld *et al.* (1976), Montanari (1974), Mackworth (1977*a*), Mackworth and Freuder (1985), McGregor (1979), Haralick *et al.* (1978), Haralick and Shapiro (1979, 1980), Haralick and Elliot (1980), Lauriere (1978), Purdom (1982, 1983), Freuder (1978, 1982), Nadel (1986, 1989, 1990*a–c*, 1991*a*), Nudel (1982, 1983*a*, *b*), Mohr and Henderson (1986), Dechter and Dechter (1987) and Dechter and Pearl (1988). As might be expected, many algorithms have been developed for solving constraint satisfaction problems. Surveys of these algorithms appear in Mackworth (1987), Nadel (1989) and Shanahan and Southwick (1989). Mathematical complexity analyses of some of these algorithms appear in Haralick and Elliot (1980) and Nudel (1983*a*).

The importance of CSP is due to the wide range of practical problems it can be used to model. Applications of the standard form of the problem or a close relative have included such diverse areas as theorem proving (Brown and Purdom 1981; Nadel 1990*b*; Van Hentenryck and Dincbas, 1986; Van Hentenryck 1989), belief maintenance (Dechter, 1987; DeKleer, 1986; Doyle 1979; Smith and Kelleher 1988), graph problems (Fowler *et al.*, 1983; McGregor, 1979; Ullman, 1976), machine vision (Barrow and Tenenbaum, 1976; Cohen and Feigenbaum, 1986; Davis and Rosenfled, 1981; Mackworth, 1977*b*; Waltz, 1975), event scheduling and general temporal reasoning (Allen, 1983; Rit, 1986; Tsang, 1987), layout and location (Eastman 1972; Navinchandra and Marks 1987*a*; Navinchandra, 1991), routing (Shanahan and Southwick, 1989), planning genetic experiments (Stefik, 1981), and micro-computer system configuration (Frayman and Mittal, 1987). A survey of some of these applications appears in (Nadel, 1990*c*). A variety of natural CSP formulations are in fact usually possible for a given real-world application. This is discussed in Nadel (1990*a*).

This paper presents our preliminary results with a system we call TRANS-FORM, that uses the CSP framework for the apparently new application of designing automatic automobile power transmissions. This is a project that we have now been pursuing for

Received 22 August 1990 and in revised form 25 June 1991.

0891–0604/91/030137 + 37$02.00/0

about 6 months in collaboration with Ford Motor Company. In particular, we have been working with the Ford Advanced Transmission Design Department in Livonia, Michigan, where our principal 'domain expert' has been Robert Roethler. Much of the following is a CSP formulation of the domain knowledge presented to us by Mr Roethler and his colleagues. Wherever possible though, we include here references to the literature in support of this domain knowledge.

Section 2 presents background material for this work. It is divided into two parts: background on the Constraint Satisfaction Problem, in Section 2.1, and background on the Transmission Design Problem, in Section 2.2. The latter discussion is rather extensive so as to provide not only a solid foundation for the current work, but also for its future extensions, such as those in Nadel (1991*b*, *c*). Our CSP formulation of the transmission design problem is given in Sections 3 and 4, with Section 3 giving the CSP variables and their domains and Section 4 giving the corresponding constraints that we use. With the CSP formulation of our problem in hand, in Section 5 we step back and look at this formulation from a more theoretical perspective, including the use of earlier mathematical results to predict the expected number of solutions for the class to which our problem belongs. The actual empirical results for our problem are given in Section 6. Expected future extensions of our TRANS-FORM system are discussed in Section 7. Some more detailed results are provided in the Appendix. A simplified version of this paper is available in Nadel and Lin (1991*a*). A discussion of the implementation of our system in Prolog appears in Nadel and Lin (1991*b*). An analysis of the complexity of solving arbitrary constraint satisfaction problems in Prolog appears in Nadel (1990*b*).

Our basic result is that we have been able to rediscover (at the kinematic level) two well-known 4R-speed[1] automatic transmissions, Axod (Ford Motor Co. 1985) and the HydraMatic 700 (Ellinger, 1983). Five[2] other 4R-speed transmissions (one of which is a 4R-speed extension of the classic 3R-speed Simpson transmission) were also found as solutions under our present formulation. It appears that most of these other five are in fact not viable. This is not surprising because we have not yet incorporated all applicable constraints at the current level of abstraction, let alone at more refined levels (increased 'depth'). Preliminary indications are that in general such extensions will be straight-forward, except for incorporation of the *topological constraints* (Nadel, 1991*c*) discussed briefly in Section 4.5 below.

Besides what we found, there is what we didn't find. Several known transmissions were overlooked. Again, this is not surprising because, in the interests of simplicity, we have made various specializations in formulating our search space. It seems that rather simple extensions to our formulation will allow all known 2-planetary transmissions to be rediscovered. Further extensions such as allowing arbitrary numbers of gearsets and speeds (increased 'breadth'), are expected to lead to the discovery of totally new transmission designs. These issues are discussed further in Sections 6 and 7.

[1] We call a transmission nR-speed if it has n forward speeds (or gear ratios or 'gears') and one reverse speed. The description n-speed, without an R, refers to a transmission with n forward speeds but no reverse.

[2] In the earlier papers Nadel and Lin (1991*a*) and the IJCAI-91 workshop version of Nadel and Lin (1991*b*), a total of 10 solutions were reported, whereas here we find only seven. This is because of a slight difference in the set of simplicity-of-switching constraints used, as described in Section 4.3.

2. Background

Our work involves the application of constraint satisfaction techniques to transmission design. In this section we present background material for both components (i) the constraint satisfaction problem and *(ii) the transmission design problem.*

2.1 THE CONSTRAINT SATISFACTION PROBLEM

Constraint satisfaction problems involve three components: *variables, values* and *constraints*. The goal is to find *all* assignments of the values to the variables such that all the constraints are simultaneously satisfied. More specifically, there is a set $Z = \{z_1 z_2 \cdots z_n\}$ of n variables z_i. Each variable takes values from an associated finite domain d_{z_i} of m_{z_i} values. There is a set $C = \{C_1 C_2 \cdots C_c\}$ of c constraints. A constraint C_j is some way of specifying for a given set $Z_j \subseteq Z$ of argument variables, which values for those variables together 'satisfy' the constraint—where values for a variable are chosen only from the corresponding domain. Thus each constraint C_j specifies a subset $T_j \subseteq D_j$ of satisfying-tuples from the Cartesian product $D_j = \times_{z_i \in Z_j} d_{z_i}$ of the domains of the constraint's arguments. We call T_j the *relation* induced by constraint C_j. A constraint may thus be specified canonically as a pair $C_j = (Z_j\ T_j)$. But this is not to imply that CSP instances must be given or solved in terms of such canonical constraints. For most CSP algorithms, a constraint may be given in arbitrary form (algebraic equations, inequalities, logical expressions, tables, procedures etc.) so long as it can be implemented as a subroutine allowing the checking of whether a given set of values satisfies the constraint. The overall Cartesian product $D = \times_{z_i \in Z} d_{z_i}$ we call the *search space* for the problem. Formally, the goal in solving a given CSP instance is to find all n-tuples in the search space D that (have projections which) satisfy all c constraints.

There are several other parameters that are important in understanding the complexity of algorithms that solve constraint satisfaction problems (Nadel, 1986, 1989, 1990*a*, 1991*b*; Nudel, 1982, 1983*a*, *b*). The *arity* of a constraint is $A_j = |Z_j|$, the number of argument variables it has. The *satisfiability* or *looseness* of a constraint is $S_j = |T_j|$, the number of value-tuples from D_j that satisfy the constraint. This number can be anywhere from 0 to the *constraint's Cartesian product size* $M_j = |D_j| = \prod_{z_i \in Z_j} m_{z_i}$. The ratio $R_j = S_j/M_j$ is the constraint *satisfiability ratio* or *looseness ratio*. It can be anywhere from 0 to 1. Higher values of S_j and R_j correspond to looser constraints. The size of the search space we denote by M. Its value is given by $M = |D| = \prod_{z_i \in Z} m_{z_i}$. Examples of these quantities for our CSP formulation of the transmission design problem will be given below, in particular in Table 4.

We will see that our version of the transmission design problem can be formulated as a constraint satisfaction problem with exactly $n = 10$ variables, two having domains of size $m_{z_i} = 9$ and eight having domains of size $m_{z_i} = 6$. There are $c = 43$ constraints, of which eight are unary, 22 are binary and 13 are quaternary. This is in fact a surprisingly small problem, of about the same order of difficulty as the 10-queens problem, which under the standard CSP formulation (Nadel, 1990*a*) has 10 variables, each of domain size 10, subject to $\binom{10}{2} = 45$ binary constraints. We will also see that from the underlying search space of over 136 million ($M = 9^2 \times 6^8$) candidate solutions, our problem has in fact only seven solutions. The 10-queens problem by contrast has 724 solutions (Nadel, 1989). In Section 5 we will use a mathematical expression based on domain sizes m_{z_i} and constraint satisfiabilities R_j to predict that for a particular class of instances to which our CSP formulation of transmission design belongs, the expected number of solutions is 6.35. This is remarkably close to the actual number of solutions, 7, found for our problem given that its search space has over 136 million candidate solutions and given that, as we will see, the expected value 6.35 is obtained as an average over a huge class of about 4.6×10^{4218} problems.

In any case, the small size of the CSP instance required to model our present version of the transmission design problem means we are not currently forced to make use of more sophisticated CSP algorithms such as Forward Checking or various arc consistency or path consistency techniques, but can make do with the relatively easy-to-understand and easy-to-implement backtracking algorithm (Nadel, 1989; Nudel, 1983*a*). And implementation is even easier if we rely on Prolog's built-in backtracking capability and use that language to solve our CSP instance as described for a range of applications in (Nadel, 1990*c*) and specifically for our present application in (Nadel and Lin, 1991*b*).

In future generalizations and extensions of our problem, the resulting CSP will become larger and we expect more sophisticated algorithms to become necessary. In this case, Prolog's built-in backtracking becomes less relevant because we will require more sophisticated approaches than straight backtracking. Such approaches can of course be implemented in Prolog too, but the intrinsic advantage of Prolog in that it implicitly carries out backtracking, will be reduced or eliminated. (A Prolog with improved methods for handling CSP built in to the interpreter would be useful. Such extensions to Prolog are in fact under development, as discussed in Van Hentenryck and Dincbas (1986) and Van Hentenryck (1989).) In any case, we do expect that, at least in conjunction with use of better CSP algorithms, a CSP formulation such as that used here will successfully scale up to handle future more complex (deeper and broader) versions of the transmission design problem. Preliminary results already indicate that this is possible in treating increases in depth to include a 'gear level' (Nadel, 1991*b*) and a 'link topology level' (Nadel, 1991*c*), as well as in treating increases in breadth to include variable numbers of gearsets and transmission speeds. The notions of search space breadth and depth are discussed more explicitly in Section 7, in connection with anticipated future extensions of our work.

For those familiar with Operations Research, CSP is closely related to linear programming, LP, and integer programming, IP. In the terminology of LP and IP, we are interested in generating what is known as the *feasible region*. But unlike in LP, CSP *per se* does not (although some extensions of CSP do) involve finding an optimum in the feasible region. Rather the whole feasible region is to be output. In LP this is not even possible in principle, because the feasible region is continuous due to the problem variables having continuous domains of values. In CSP it *is* possible since all domains are assumed discrete and finite. Integer programming is similar in having discrete domains (of integers), although the domains are not necessarily finite. So for IP too, unlike for CSP, a listing of the whole feasible region is not generally possible. CSP also differs from IP in that variables are allowed to take values of arbitrary type (not necessarily just integers nor even numbers) and constraints accordingly may be of arbitrary type rather than just linear inequalities. Even though the feasible region is finite in our CSP formulation of transmission design, and may thus be fully listed explicitly, it is nevertheless natural (as in many other applications of CSP) to seek an optimum in the feasible region. However, the preliminary formulation described here does not consider optimization issues. Later formulations probably will. This is discussed further in Section 7.

2.2 THE TRANSMISSION DESIGN PROBLEM

2.2.1 Automatic transmissions

An automobile must deliver *torque* or 'turning power' from the engine, via what is called the *drive train,* to the driving wheels. In rear wheel drive vehicles the drive train consists of a *transmission,* a *drive shaft* and a *rear axle assembly,* whereas in front wheel drive vehicles all these components are integrated in a *transaxle* assembly.

In either case, the basic component of the drive train is the transmission. It acts as a torque multiplying (and dividing) device for adjusting the amount and direction of torque delivered from the engine to the drive wheels under varying engine operating conditions and driving conditions. The torque multiplication factor is known as the transmission's *gear ratio.* A *manual transmission* requires the driver to control the gear ratio manually, via *countershaft gears.* These are sets of gears on two parallel axes. In their simplest form, change of gear ratio is achieved by sliding the axes relative to each other so as to cause different pairs of gears to engage. On the other hand, automatic transmissions, with which we concern ourselves here, determine the gear ratio automatically, and usually use *planetary gearsets* rather than countershaft gears. As will be described in more detail below, a planetary gearset is a combination of *sun, ring* and *planet* gears arranged somewhat like a miniature solar system. Good introductions to automatic transmissions and planetary gearsets are found in Ellinger (1983), Husselbee (1986) and Juvinall (1983). A more advanced treatment can be found in Müller (1982). Compared to manual transmissions:

- Automatic transmissions have the obvious advantage of being easier to use.
- As mentioned in Husselbee (1986) automatic transmissions also have several other advantages. In (at least the earlier versions of) manual countershaft transmissions, un-meshing and re-meshing of gears is required in order to change gear ratios. This leads to the possibility of gear teeth damage due to 'crunching' the gears during partial engagement or because of improper shifting techniques. In automatic transmissions the ratio is changed simply by changing which components of the planetary gearsets are 'input', 'output', 'braked' and 'linked' (as we will see below), thus allowing the gears of an automatic transmission to stay in constant mesh. Gear crunching is thus eliminated.
- The above constant-mesh feature of automatic transmissions allows quick ratio changes, avoiding the loss of torque flow which may be experienced during gear changing in manual transmissions.
- Automatic transmissions are stronger for a given size, since they distribute the torque load over more gears.
- Automatic transmissions are more compact due to the coaxial nature of the planetary gearsets.

An automatic transmission may be considered as made up of three interrelated subsystems: a *mechanical,* a *hydraulic* and an *electronic* subsystem, the latter two subsystems being for the purpose of controlling the former. Our current work concentrates only on the design of the mechanical subsystem of automatic transmissions. We expect to extend this to an integrated treatment of the design of all three subsystems in the future.

2.2.2 Planetary gearsets

Most automatic automobile transmissions are made from various combinations of various types of *planetary gearsets.* We assume here that only *simple,* as opposed to *compound,* planetary gearsets are used. A description of compound gearsets is beyond the scope of this paper. See for example Husselbee (1986) or Lynwander (1983). Figure 1 shows an example of a simple planetary gearset with four planets. (Figure 2 shows an abstraction with three planets.) The planets *p* are attached to an arm *a* which can rotate with the planets about a central sun gear *s*. In doing so, the motion of the planet gears is analogous to that of planets in a solar system, which is of course the reason for the name planetary gearset. Each planet rotates about its own center as the group of planets on the arm revolves about the sun. At the perimeter of the gearset is a ring gear *r* whose teeth are on the inside so as to mesh with the planets. Normal gears like the sun and planet gears, where the teeth point outwards,

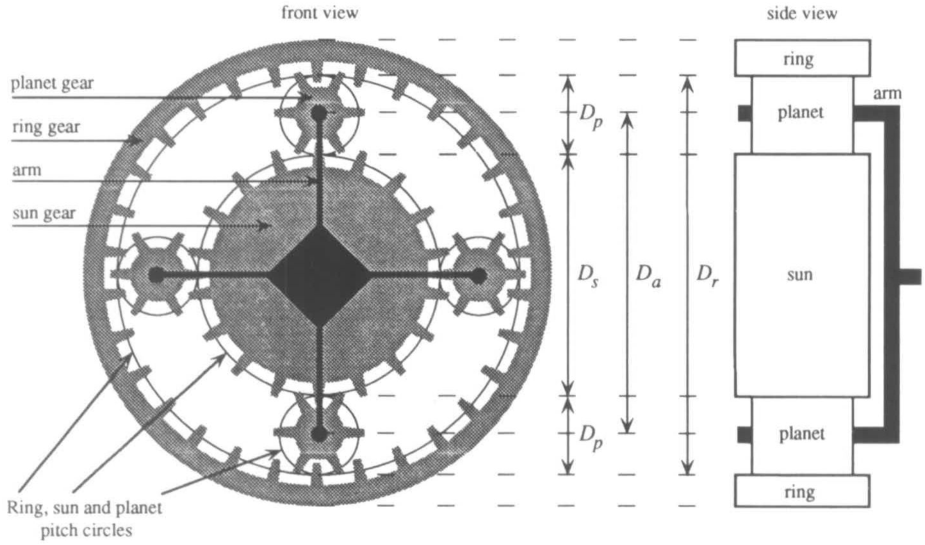


FIGURE 1. A planetary gearset with four planet gears. There are $T_s = 18$ teeth on the sun, $T_p = 6$ on each planet and $T_r = 30$ on the ring. Note that the corresponding pitch circle diameters have the same ratios as the teeth numbers

are called *external gears*. Gears like the ring, where the teeth point inwards are called *internal* or *annulus gears*.

Several different circles characterize an individual gear (Juvinall, 1983). There is the *dedendum circle*, the *base circle*, the *pitch circle* and the *addendum circle*. The dedendum circle joins the bottoms of the 'valleys' between teeth. The addendum circle joins the 'peaks' of the teeth. The other circles fall in between. For the purposes of kinematic analysis we may model a pair of meshing gears by an equivalent pair of cylinders pressed together which can turn each other via friction without slippage. The pitch circle of a gear may be thought of as the cross section of the corresponding cylinder in such a pair of kinematically equivalent cylinders. Pitch circles are shown for the component gears in Figure 1. It is the diameter D of a gear's pitch circle—the gear's *pitch diameter*—that is the relevant diameter for kinematic analysis. For instance, two meshing gears with fixed centers have angular velocities inversely proportional to their pitch diameters.

The constant of proportionality between the number of teeth T on a given gear and its pitch diameter is called the gear's *diametral pitch*, which we denote as α. Thus we have $T = \alpha D$. It should be clear that for two gears to be able to mesh they must have the same number of teeth per unit of their circumference, and hence per unit of their diameter. Thus for the sun (s), planet (p) and ring (r) gears of a *single* planetary gearset, we have

$$T_s = \alpha D_s, \qquad T_p = \alpha D_p, \qquad T_r = \alpha D_r, \quad \text{all for the } \textit{same} \text{ constant } \alpha \tag{1}$$

The value of α is thus fixed for all gears in a given gearset, so that α is as much a parameter of a gearset as it is of a gear in a gearset. However, different gearsets may of course have different values of α.

An important parameter for our purposes will be the ratio of the number of teeth T_r on the ring to the number of teeth T_s on the sun in a gearset,

$$\beta = T_r/T_s. \tag{2}$$

In the transmission of Figure 1 for example, the value of this parameter is $\beta = 30/18 = 1.666$. We will see below (in connection with Figure 7) that the assumptions we make in the current formulation of the transmission design problem allow gearsets corresponding to $7/5 \leq \beta \leq 7/2$. For discussion purposes we will often use the nominal value of $\beta = 2$ within this range (as for example when giving the approximate numerical values for gear ratios ρ in Tables 1, 2 and 3 below). Note also that the following relationship (obvious from Figure 1) holds between the pitch circle diameters of component gears in a planetary gearset:

$$D_r = D_s + 2D_p. \tag{3}$$

2.2.3 *Equations of motion for a planetary gearset*

A basic requirement in meshing gears (whether in a planetary gearset or not) is that the tangential

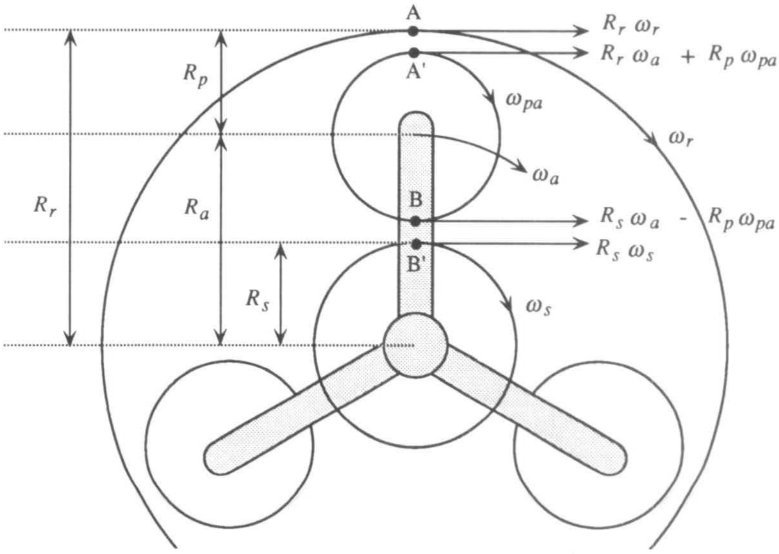


FIGURE 2. Linear and angular velocities in a planetary gearset

velocities of the points of contact must be equal. From this we can derive the equations of motion for the components of a planetary gearset. Figure 2 provides the notation we will use.[3] The circles shown there are the pitch circles for the sun, planet and ring gears of the gearset. As such, points A and A′ should touch, as should points B and B′. The circles are shown separated only to help in the following discussion where we need to distinguish between touching points on different meshing gears.

Let ω_s, ω_a, ω_r and ω_p denote the angular velocity (in radians per second) of the sun, arm, ring and planet of a gearset. (All planets have the same angular velocity.) These angular velocities are relative to the frame or housing of the gearset and may be considered as *absolute* angular velocities. On the other hand, let ω_{pa} be the *relative* angular velocity of the planet with respect to the arm of the gearset, so that

$$\omega_p = \omega_a + \omega_{pa}. \tag{4}$$

The arc length x spanned by a radius of length r rotating through θ radians is $x = r\theta$. Dividing both sides by the unit of time, we have that the speed v of the tip of a radius of length r rotating at uniform angular velocity ω is $v = r\omega$. Thus relative to the arm, point B on the planet has velocity $v_{pa}^B = R_p\omega_{pa}$. Similarly, the arm itself at the point corresponding to B is moving at velocity $v_a^B = R_s\omega_a$. Using for linear velocities the counterpart of the angular velocity relation (4), we may express the absolute velocity of point B as $v_p^B = v_a^B + v_{pa}^B = R_s\omega_a - R_p\omega_{pa}$, where the negative is due to the opposite senses of the two component velocities. Point B′ on the sun has absolute velocity $v_s^{B'} = R_s\omega_s$. Since B′ and B denote meshing points, their velocities must be equal, so that $v_s^{B'} = v_p^B$ and hence $R_s\omega_s = R_s\omega_a - R_p\omega_{pa}$. A similar analysis regarding the speeds of points A and A′ leads to $R_r\omega_r = R_r\omega_a + R_p\omega_{pa}$. Using (1) we may re-express the latter two equations in terms of teeth numbers $T_i = \alpha D_i = 2\alpha R_i$, the same α for each gear $i = s, p, r$. We thus obtain

$$T_s\omega_s - T_s\omega_a + T_p\omega_{pa} = 0 \tag{5}$$

and

$$T_r\omega_a - T_r\omega_r + T_p\omega_{pa} = 0. \tag{6}$$

From these we have

$$T_s\omega_s + (-T_r - T_s)\omega_a + T_r\omega_r = 0 \tag{7}$$

and

$$T_s\omega_s + (-T_p - T_s)\omega_a + T_p\omega_p = 0 \tag{8}$$

in terms of only the absolute angular velocities ω_s, ω_a, ω_r and ω_p, without the relative velocity ω_{pa}. Equation (7) is obtained by subtracting (6) from (5), and (8) is obtained from (5) by using (4) to eliminate ω_{pa}.

2.2.4 *Kinematic states of a planetary gearset*

A planetary gearset provides the basis for a remarkably flexible mechanism for changing angular

[3] Note that the arrows in Figure 2 are drawn so as to point in the corresponding positive directions, which in our convention is rightwards for linear velocities and clockwise for angular velocities. The arrows are not intended to show the actual direction of motion of the corresponding parts. The actual directions of motion are obtained from the signs of the corresponding values when the equations of motion are solved. If the motion is in the opposite sense to that shown by the arrow then the velocity value will be negative.

velocity (and torque). There are many combinations of motion possible, all of which may be understood in terms of the equations of motion (7) and (8). However for our purposes the cases of interest are only those where one of the component parts, sun s, arm a or ring r is linked to an *input* torque (corresponding to the vehicle's motor), one of the other two parts is held fixed or *grounded* or *braked* and the third part is used to drive an *output* (corresponding to the vehicle's wheels). Note that this excludes the possibility of the planets being linked to input, being braked or being linked to output. In principal this is also possible, but in practice it is usually not of interest due to the difficulty of building a housing for the planetary gearset which allows these operations on planets *per se*, since the latter not only rotate on their axes but revolve about the sun while doing so. On the other hand, the sun, arm and ring simply rotate about their common axis. It is thus relatively straightforward to build a housing which allows co-axial input to, output from or braking of, the sun, arm or ring, but not the planets. Thus in automobile transmissions the planets of a gearset are used just as intermediaries to transfer motion between the other components.[4]

In choosing a gearset component s, a or r to use for input there are three possibilities, for each of which there are two possibilities in choosing the remaining component to hold fixed, for each of which there is only one possibility for choosing the remaining component to use as output. The corresponding $3 \times 2 \times 1 = 6$ configurations are shown as cases 1 to 6 in Table 1. (Cases 7 and 8 shown there will be discussed later.) The diagrams used to denote planetary gearsets in Table 1 and in other tables below, are intended as schematics of the *top half* of the side view of a planetary gearset such as shown at the right in Figure 1. Note however that, unlike in such a top-half gearset side view, the middle squares of our schematic diagrams do not denote a planet *per se*, but rather the arm which connects to planets. The schematic diagrams use inwards and outwards arrows to denote the input and output components respectively, and use gray shading to denote the braked component.

As seen in Table 1, each of the states 1 to 6 results in a different gear ratio $\rho = \omega_{Out}/\omega_{In}$, the ratio of the angular velocity of the corresponding output component *Out* to that of the input component *In*. These ratios may be obtained by use of equation (7) in conjunction with two other simple equations characterizing the input and the braked components. Equation (8) is superfluous for our purposes since we are not interested in the angular velocity ω_p of the planets.[5] We are interested in only the three unknowns ω_s, ω_a and ω_r, the angular velocity of the sun, arm and ring respectively. We may obtain three linear equation in these three unknowns as follows. The first equation is (7) as mentioned (where for the versions in Table 1 we have divided through by T_s to give the coefficients in terms of $\beta = T_r/T_s$). It characterizes the overall planetary gearset itself, not any particular state, so it is part of the characterization of each of the states in Table 1. The second equation characterizes the braked component. In state 1 of the table for example this is the arm. Since the arm is braked it has zero angular velocity, and we may thus add the corresponding equation $\omega_a = 0$. The third equation characterizes the input component. In state 1 this is the sun gear. We can then write $\omega_s = c$, where c denotes the input velocity of the sun. Since we are not interested in the case of any specific input velocity, but rather in the *ratio* of output velocity to input for arbitrary input velocity, we may as well use $c = 1$. Using $\omega_{In} = c = 1$ is convenient as it allows a gear ratio to be obtained as $\rho = \omega_{Out}/\omega_{In} = \omega_{Out}/1 = \omega_{Out}$.[6] We thus have for state 1 the three equations

$$\begin{aligned} \omega_s + (-\beta - 1)\omega_a + \beta\omega_r &= 0 \\ \omega_a &= 0 \\ \omega_s &= 1 \end{aligned}$$

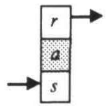
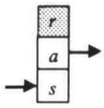
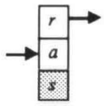
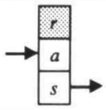
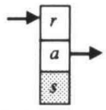
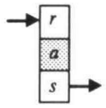
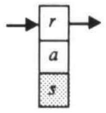
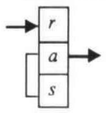
It is these equations written in matrix form which appear in the 'Kinematic Equations' column for state 1 in Table 1. The other states' equation sets are obtained similarly. Solving these equations for the angular velocity of the respective output components (remember $\rho = \omega_{Out}$) leads to the expressions shown in the table for gear ratio $\rho(\beta)$ as a function of β.

[4] An interesting contrast occurs however in the transmission of say a mixing machine. In such a case the rotation plus revolution of the planets is a very natural motion for stirring, which it is not even necessary to build an actual hardware link to capture. The motion may simply be passed on to the stirred mixture directly via paddles connected to the planets' centers. In such a case planets could be, and are, conveniently used for output.

[5] However, in a graphics program that we have written to animate the motion in planetary gearsets, equation (8) did come in handy because there it was necessary to solve for the angular velocity of the planets also so as to allow a complete animation. In our later, more-detailed design work as well equation (8) will become relevant since a full design must satisfy constraints on the maximum speed at which each gearset component, including the planets, may turn so as not to overload its bearings. This type of constraint is not considered in our current formulation.

[6] Leaving c arbitrary would give the same ratio of output to input velocity, but less conveniently. The output velocity would simply be obtained as a function of c, and c in the numerator and denominator of $\rho = \omega_{Out}/\omega_{In}$ would cancel in obtaining the ratio.

TABLE 1. Gear ratios achievable using a single planetary gearset. States 7 and 8 represent two different approaches to achieving direct drive

State	In	Braked	Out	Diagram	Kinematic Equations	Gear Ratio $\rho = \frac{\omega_{Out}}{\omega_{In}}$		Ratio Type
1	s	a	r		$\begin{bmatrix} 1 & -\beta-1 & \beta \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \omega_s \\ \omega_a \\ \omega_r \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$	$-\frac{1}{\beta}$	$\approx -1/2$	reverse underdrive
2	s	r	a		$\begin{bmatrix} 1 & -\beta-1 & \beta \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \omega_s \\ \omega_a \\ \omega_r \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$	$\frac{1}{1+\beta}$	$\approx 1/3$	underdrive
3*	a	s	r		$\begin{bmatrix} 1 & -\beta-1 & \beta \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \omega_s \\ \omega_a \\ \omega_r \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$	$\frac{1+\beta}{\beta}$	$\approx 3/2$	overdrive
4	a	r	s		$\begin{bmatrix} 1 & -\beta-1 & \beta \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \omega_s \\ \omega_a \\ \omega_r \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$	$1+\beta$	≈ 3	overdrive
5*	r	s	a		$\begin{bmatrix} 1 & -\beta-1 & \beta \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \omega_s \\ \omega_a \\ \omega_r \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$	$\frac{\beta}{1+\beta}$	$\approx 2/3$	underdrive
6	r	a	s		$\begin{bmatrix} 1 & -\beta-1 & \beta \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \omega_s \\ \omega_a \\ \omega_r \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$	$-\beta$	≈ -2	reverse overdrive
7	r	s	r		As for state 5: $\begin{bmatrix} 1 & -\beta-1 & \beta \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \omega_s \\ \omega_a \\ \omega_r \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$	1		direct drive
8	r	–	a		$\begin{bmatrix} 1 & -\beta-1 & \beta \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \omega_s \\ \omega_a \\ \omega_r \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$	1		direct drive

States 3 and 5 (*) are respectively the overdrive and underdrive states of the classic 3-speed Sturmey-Archer bicycle transmission.

Note that unlike the choice of input and braked part, the choice of output part does not effect a state's set of equations *per se*. But it does of course effect the gear ratio for a state by determining which solution, ω_s, ω_a or ω_r, of the equation set will be used to obtain $\rho = \omega_{Out}$.

Our algorithm will search over the feasible kinematic states for a transmission and for each of these will dynamically generate a corresponding set of simultaneous equations like those in Table 1 (but for a transmission made of two linked gearsets, as will be discussed). Each equation set generated is solved symbolically to obtain the corresponding symbolic expression for ρ, analogous to those in Table 1. Substituting a numerical value for β would of course allow us to obtain numerical ρ values for checking say if an acceptable ratio value is achieved for a given speed. Besides giving the gear ratios $\rho(\beta)$ symbolically, Table 1 also gives corresponding example numerical values using the nominal value of $\beta = 2$. This is within what we will see is our allowed range of $7/5 \leq \beta \leq 7/2$. At this stage of our formulation however, our algorithm does not consider numerical values for β or ρ, but reasons directly with the symbolic ρ expressions, as will be described below in connection with (14).

Negative values for ρ correspond to *reverse* gear ratios: the direction of rotation between input and output is reversed. Values of ρ over 1 correspond to *overdrive* gear ratios: the angular velocity of the output is greater than that of the input. Values of ρ under 1 correspond to *underdrive* gear ratios: the angular velocity of the output is less than that of the input. A value of ρ equal to 1 corresponds to a *direct* *drive* gear ratio: the angular velocity of the output equals that of the input. We see from Table 1 that a single planetary gearset is capable of delivering quite a range of gear ratios: two reverse ratios, two underdrives, two overdrives and a direct drive, the latter being achievable in a variety of ways. States 7 and 8 are representative of two different approaches for achieving direct drive.

The first approach, represented by state 7, is to use for the output the same component as is used for input. Note that the three equations characterizing a state do not reflect the choice of output component. They only reflect the choice of input and braked components. Thus for example, since state 7 has the same input and braked components as state 5 in the table, the corresponding equations are the same. The only difference between the states is in the choice of output, which in terms of the equations is reflected only as a difference in which variable is being solved for. In state 7 for example we are solving for ω_r, rather than ω_a in state 5. Of course situations like state 7 are degenerate cases, since the ratio of output to input velocity is the ratio of the output velocity to itself.

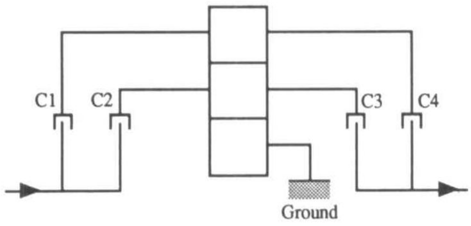
The second approach to obtaining direct drive in a planetary gearset is represented by state 8 of Table 1. Here we are allowing an added link between the arm and the sun so that they rotate in unison; formally this link is represented by the equation $\omega_s = \omega_a$ or $\omega_s - \omega_a = 0$. In this added-link approach, there is no need for a braked component in order to get a well defined system, since we already have three equations in three unknowns given that one of the components is used for input. This can be seen for our example in the table's equations column for state 8. One equation characterizes the gearset *per se,* one the link and one the input. Note there is no shaded square for a braked component in the state 8 diagram.

A little thought shows that when there is a link between any two of the sun, arm and ring components, the gear ratio will always be 1, no matter what are the two linked components and no matter what component is used for output. Since there are so many ways of achieving direct drive, via either of the approaches represented by states 7 or 8, we have not bothered to require a direct drive state in the transmissions we are currently designing. We are in effect assuming that any solution which satisfies the other constraints on gear ratios etc. can be extended to have a compatible direct drive state. Later formulations will not make this assumption, but it is a useful and relatively innocuous simplification at this stage.

Note that when we do, in later work, consider explicit direct drive states to incorporate into our designs, the approach represented by state 8 of Table 1, where an extra link is used, has several advantages compared to that represented by state 7, where the output component is simply set to the input component. When the added-link approach is used the various components, including the planets, all become locked together and rotate in synchrony, so that $\omega_s = \omega_a = \omega_r = \omega_p$. This can be seen from solving the corresponding kinematic equations such as those for state 8 in Table 1. [Equation (8) is needed in verifying the above equality for ω_p.] However, when the input-equals-output method is used to obtain direct drive, the component that is not input/output or fixed must rotate relative to the other two components, again as can be seen from solving the corresponding kinematic equations such as those for state 7 in Table 1. The added-link approach represented by state 8 thus locks the parts to rotate synchronously whereas the approach of state 7 allows relative motion. The former approach is preferable since the lack of relative motion of the parts (i) reduces gear noise, (ii) reduces gear wear and (iii) improves the power transfer efficiency due to heat, noise and general friction losses being reduced.

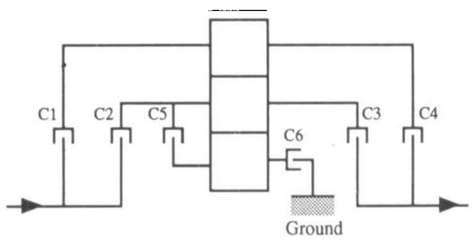
2.2.5 *One-planetary transmissions*

Though, as we have seen, a single planetary may provide quite a range of different gear ratios, it is not in itself a *transmission.* A transmission must also provide the hardware for switching between the desired kinematic states provided by one or more planetary gearsets. It is not in fact topologically possible to build a switching mechanism capable of using all seven ratios provided by a single planetary. Figures 3 and 4 are examples of single-planetary transmissions that *are* topologically realizable. Note the characteristic addition to the gearset *per se,* of a network of clutched links and brakes for reallocating which element is input, output and braked. The 3/4 squares ⊓ shown in these transmission 'stick diagrams' denote clutches, which may be thought of as simply switches for connecting the two parts of the corresponding path (shown by a line) to allow flow of torque along it. The tables at the right of the transmission stick diagrams include columns denoting which clutches are on in achieving the transmission's various subsumed states/speeds/ratios. Corresponding to each clutching pattern is an 'Effective Transmission' diagram denoting the effective torque pathways achieved by that clutching configuration. And each of these is shown with the corresponding designation of the input braked and output component, and the



Speed	State	Clutches				Effective Transmission	In	Braked	Out	Ratio ρ	Ratio Type
		C1	C2	C3	C4						
1	5	On		On			r	s	a	$\frac{\beta}{1+\beta}$	under drive
2	7	On			On		r	s	r	1	direct drive
3	3		On		On		a	s	r	$\frac{1+\beta}{\beta}$	over drive

FIGURE 3. A Sturmey-Archer-like 3-speed transmission based on a single planetary gearset, using a common part for input and output to obtain direct drive



Speed	State	Clutches						Effective Transmission	In	Braked	Out	Ratio ρ	Ratio Type
		C1	C2	C3	C4	C5	C6						
1	5	On		On			On		r	s	a	$\frac{\beta}{1+\beta}$	under drive
2	8	On		On		On			r	—	a	1	direct drive
3	3		On		On		On		a	s	r	$\frac{1+\beta}{\beta}$	over drive

FIGURE 4. Another Sturmey-Archer-like 3-speed transmission based on a single planetary gearset, using a clutched (or soft) link to obtain direct drive

corresponding gear ratio and ratio type, in a manner similar to that of Table 1.

We see that both transmissions incorporate states 3 and 5 of Table 1 for their overdrive and underdrive states respectively. The transmissions differ in which state they use for direct drive. This difference imposes the need for a different switching network, as seen. Nevertheless, the two transmissions are very similar. We call two transmissions *functionally equivalent* if they incorporate the same set of gear ratios. We say they are *kinematically equivalent* if in addition the same kinematic states (input, output and braked components) are used for corresponding ratios. The transmissions in Figures 3 and 4 are functionally equivalent but not kinematically equivalent. They are in fact functionally equivalent to the classic Sturmey-Archer 3-speed bicycle transmission (Hadland, 1987), and kinematically equivalent to it, and to each other, in all but the direct drive state. They are 'kinematically equivalent modulo direct drive'.

As mentioned earlier, the approach of state 8 of Table 1 is preferable to that of state 7 in achieving direct drive, since it reduces gear noise, gear wear and improves power transfer efficiency for that gear ratio. In this sense the transmission of Figure 4 is preferable to that of Figure 3. On the other hand, that of Figure 3 is preferable in having less clutches, and thus being cheaper to build and easier to maintain. But the transmission of Figure 4 may be simplified (to a version that is kinematically equivalent modulo direct drive) while still retaining its advantage of using the added-link approach to achieving direct drive. One could for instance just remove clutch C5 and its path completely, and obtain direct drive by turning on clutches C1, C2 and C3 simultaneously, rather than C1, C3 and C5 as now. This simplifies the transmission from six clutches to five, but the transmission of Figure 3 is still preferable in having even less (four) clutches, although it is worse in noise, wear and efficiency as mentioned.

The above issues of simplification of the switching network in a transmission, the reduction of noise, wear and inefficiency, and the trade off between these is an example of the multi-criteria optimization which needs to be addressed in the full version of the transmission design problem. We do not pursue any optimization in our present formulation. In fact, the particular issues of complexity of switching network, gear noise, gear wear and power transfer efficiency do not even arise at the level of abstraction at which we are currently treating transmission design. In particular, we ignore the switching network, the specific nature of the direct drive state, and the number and shape of gear teeth (the latter two parameters have an effect on gear noise, wear and transmission efficiency when there is relative gear motion whether in direct drive or other states).

As indicated in Figures 3 and 4, the various kinematic states of a transmission are naturally ranked according to the magnitude of their gear ratios and are said to be of successive 'speeds'. The transmissions in Figures 3 and 4 for example have three speeds, from speed 1 for the underdrive state, through speed 2 for the direct drive state, to speed 3 for the overdrive state. In automobile transmissions we will also want a single reverse speed, which we denote as speed 0. There were no reverse speeds in our examples of Figures 3 and 4 because the Sturmey-Archer transmission on which they are based is a bicycle transmission, and people don't usually want to ride backwards on bicycles. In automobiles of course, a reverse gear is desirable. Some vehicles such as trucks and tractors may even have several negative gear ratios of different magnitudes, and hence several reverse speeds or reverse 'gears'. As here, the term *gear* is often used for *speed* or *gear ratio*. This is of course a different sense of the word than when talking about the physical gears of say a planetary gearset. Context should clarify the sense intended.

The transmissions of Figures 3 and 4 provide examples of another issue relevant in transmission design, *simplicity of switching*. This refers not to the above-mentioned simplicity of the switching network *per se*, but rather to the simplicity of the changes in clutching pattern needed to go between successive speeds of the transmission. We leave the discussion of simplicity of switching to Section 4.3 below. Suffice it to say here that the transmission of Figure 3 is preferable to that of Figure 4 (and also to its network-simplified version discussed above) in terms of simplicity of switching (as well as in terms of simplicity of the switching network). In fact, as we will see, neither version of the Figure 4 transmission satisfies even the minimal simplicity-of-switching constraints which we are currently using.

Note that in building a transmission which incorporates various states it is desirable to have the same member (sun, arm or ring) acting as output in each state. (Actually, we will see in Section 3.3 that it is undesirable to allow output from a sun gear.) This is because the output is where relatively large torques can occur in an automobile transmission; an automobile's engine always rotates at relatively high speed but low torque, whereas when climbing a steep gradient say, the wheels may rotate slowly but with high torque. To allow for the inevitable situations when the output does involve high torque, it is best to have the output component linked to the wheels by a relatively sturdy permanent or 'hard' link rather than by a relatively weak temporary or 'soft' link achieved via clutching. Thus one prefers to achieve the necessary range of transmission ratios by using clutches to vary which are the input and braked components rather than to vary which member is the output component. In the transmissions of Figures 3 and 4 this was not actually the case, but again this is because the Sturmey-Archer transmission, on which they were based, is intended for bicycles where the required torque is never great anyway. For automobile transmissions on the other hand torques are much larger so that having a fixed output component becomes important.

But from Table 1 we see that with no fixed choice of output member does one achieve a useful distribution of ratios (even counting the direct drive ratio of 1). If the sun is the fixed output, we get no underdrive ratio. If the arm is the fixed output, we get no reverse or overdrive. If the ring is the fixed output, we get no (forward) underdrive. To provide a wider choice of states so that the output may be fixed without losing needed ratios and/or so that topological incompatibilities may be avoided in combining needed ratios, we need a more flexible transmission than can be built from a single planetary gearset. Transmissions made of two or more coupled gearsets give us this extra flexibility.

2.2.6 Kinematic states of two linked planetary gearsets

Coupling together two or more planetary gearsets provides a wider range of achievable kinematic states and gear ratios and hence more flexibility in satisfying constraints and/or in achieving a better optimum, in designing a transmission. In automobiles, transmissions with two and three gearsets are common. For our present initial study we restrict ourselves to two gearsets, which we distinguish as gearset 1 and gearset 2. Their relevant component sets are respectively $Parts_1 = \{s1, a1, r1\}$ and $Parts_2 = \{s2, a2, r2\}$, where *si, ai* and *ri* of course denote respectively the sun, arm and ring of gearset i. Note the absence of planets in these sets, as discussed in connection with footnote 4 p. 143. For the purposes of ordering our computer search we assume that these six components of our two gearsets have the following (purely arbitrary) relative ordering

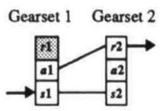
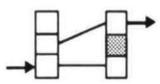
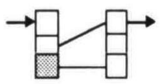
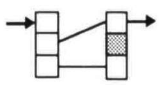
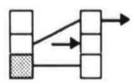
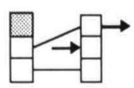
$$s1 < a1 < r1 < s2 < a2 < r2. \tag{9}$$

In linking two gearsets together it is possible to use various numbers of hard (or permanent or non-clutched) links. We currently restrict ourselves to the case of exactly two hard links. Such a pair of hard links may be configured in many ways. We will see later (in connection with Figure 8) that of these, there are only 18 pairs that satisfy the constraints of our domain. Tables 2 and 3 correspond to two of these 18 cases. Table 2 is for the two hard links $L1 = (s1, s2)$ and $L2 = (a1, r2)$, and Table 3 is for the two hard links $L1 = (a1, r2)$ and $L2 = (r1, a2)$, where $Li = (x, y)$ denotes a link between components x and y.

For each of the 18 acceptable ways of connecting two gearsets with two hard links there are many more non-direct-drive gear ratios achievable than only the six seen in Table 1 to be available with a single gearset. With two gearsets, there are in fact six non-direct-drive ratios for *each* choice of output component, as explained below in Section 4.2 (after C_7^i). Tables 2 and 3 show states corresponding to the six such ratios for the fixed output of $Out = r2$. These results are extended in the appendix, where the corresponding six states are given for each of the 18 legal hard link pairs.

Tables 2 and 3 are analogs of the earlier Table 1, and give essentially the same type of information. Note however the difference in the kinematic equations that need to be solved in obtaining the gear

TABLE 2. Gear ratios achievable using two planetary gearsets with two links $L1 = (s1, s2)$ and $L2 = (a1, r2)$, and fixed output $Out = r2$ (or $a1$)

State	In	Braked	Out	Diagram	Kinematic Equations	Gear Ratio $\rho = \frac{\omega_{Out}}{\omega_{In}}$	Ratio Type
1	s1 (s2)	r1	r2 (a1)		$\begin{bmatrix} 1 & -\beta_1-1 & \beta_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\beta_2-1 & \beta_2 \\ 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \omega_{s1} \\ \omega_{a1} \\ \omega_{r1} \\ \omega_{s2} \\ \omega_{a2} \\ \omega_{r2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	$\frac{1}{1+\beta_1} \approx 1/3$	under drive
2*	s1 (s2)	a2	r2 (a1)		$\begin{bmatrix} 1 & -\beta_1-1 & \beta_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\beta_2-1 & \beta_2 \\ 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \omega_{s1} \\ \omega_{a1} \\ \omega_{r1} \\ \omega_{s2} \\ \omega_{a2} \\ \omega_{r2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	$-\frac{1}{\beta_2} \approx -1/2$	reverse under drive
3*	r1	s1 (s2)	r2 (a1)		$\begin{bmatrix} 1 & -\beta_1-1 & \beta_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\beta_2-1 & \beta_2 \\ 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \omega_{s1} \\ \omega_{a1} \\ \omega_{r1} \\ \omega_{s2} \\ \omega_{a2} \\ \omega_{r2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	$\frac{\beta_1}{1+\beta_1} \approx 2/3$	under drive
4*	r1	a2	r2 (a1)		$\begin{bmatrix} 1 & -\beta_1-1 & \beta_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\beta_2-1 & \beta_2 \\ 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \omega_{s1} \\ \omega_{a1} \\ \omega_{r1} \\ \omega_{s2} \\ \omega_{a2} \\ \omega_{r2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	$\frac{\beta_1}{1+\beta_1+\beta_2} \approx 2/5$	under drive
5	a2	s1 (s2)	r2 (a1)		$\begin{bmatrix} 1 & -\beta_1-1 & \beta_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\beta_2-1 & \beta_2 \\ 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \omega_{s1} \\ \omega_{a1} \\ \omega_{r1} \\ \omega_{s2} \\ \omega_{a2} \\ \omega_{r2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	$\frac{1+\beta_2}{\beta_2} \approx 3/2$	over drive
6	a2	r1	r2 (a1)		$\begin{bmatrix} 1 & -\beta_1-1 & \beta_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\beta_2-1 & \beta_2 \\ 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \omega_{s1} \\ \omega_{a1} \\ \omega_{r1} \\ \omega_{s2} \\ \omega_{a2} \\ \omega_{r2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	$\frac{1+\beta_2}{1+\beta_1+\beta_2} \approx 3/5$	under drive

Unlike the single gearset case in Table 1, here reverse, underdrive and overdrive are all achievable even with the output fixed. States 2, 3 and 4 (*) are respectively the reverse, high underdrive, and low underdrive states of the classic Simpson 3R-speed automobile transmission.

ratio for a state. In the one-gearset case of Table 1 there were only three equations in three unknowns, whereas in the two-gearset case of Tables 2 and 3 there are six equations in six unknowns. The doubling of the number of variables is because now there is a sun, arm and ring gear for each of two gearsets, rather than just for a single gearset, and for each of these six parts we need to solve for its angular velocity. That is, we now have variables ω_{si}, ω_{ai} and ω_{ri} for both $i = 1$ and $i = 2$, rather than just ω_s, ω_a and ω_r.

Since there are six variables now, we require six equations for a well-defined system. As before we use equation (7) to characterize a gearset, but since there are now two gearsets we need a version of (7) for each gearset—each version of course being in terms of the corresponding gearset's three variables. Thus we have

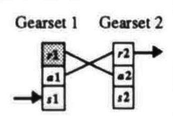
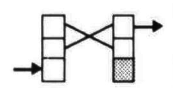
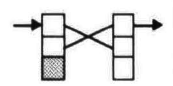
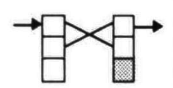
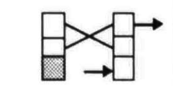
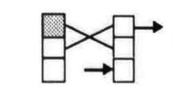
$$T_{s1}\omega_{s1} + (-T_{r1} - T_{s1})\omega_{a1} + T_{r1}\omega_{r1} = 0.$$
$$T_{s2}\omega_{s2} + (-T_{r2} - T_{s2})\omega_{a2} + T_{r2}\omega_{r2} = 0.$$

These two equations apply, and appear in the kinematic equations column, for every state in Tables 2 and 3 since all the states involve the corresponding two gearsets. As in Table 1, in Tables 2 and 3 these two equations are divided through by T_{s1} and T_{s2} respectively so that the coefficients are in terms of the ratios

$$\beta_1 = T_{r1}/T_{s1} \quad \text{and} \quad \beta_2 = T_{r2}/T_{s2}, \tag{10}$$

the gearset 1 and 2 counterparts of β, the single gearset ring-to-sun tooth-ratio defined in equation (2). The remaining four equations for a two-gearset state reflect its unique combination of hard link pair, input part and braked part. Table 2 is for the case of hard links $L1 = (s1, s2)$ and $L2 = (a1, r2)$. The first link restricts the angular velocity of $s1$ to equal that of $s2$, which contributes equation $\omega_{s1} = \omega_{s2}$ or equivalently $\omega_{s1} - \omega_{s2} = 0$. Similarly the second hard link contributes $\omega_{a1} - \omega_{r2} = 0$. All states of Table 2 have these

TABLE 3. Gear ratios achievable using two planetary gearsets with two links $L1 = (a1, r2)$ and $L2 = (r1, a2)$, and fixed output $Out = r2$ (or $a1$)

State	In	Braked	Out	Diagram	Kinematic Equations	Gear Ratio $\rho = \frac{\omega_{Out}}{\omega_{In}}$	Ratio Type
1†	$s1$	$r1$ ($a2$)	$r2$ ($a1$)		$\begin{bmatrix} 1 & -\beta_1-1 & \beta_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\beta_2-1 & \beta_2 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \omega_{s1} \\ \omega_{a1} \\ \omega_{r1} \\ \omega_{s2} \\ \omega_{a2} \\ \omega_{r2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	$\frac{1}{1+\beta_1} \approx 1/3$	under drive
2†	$s1$	$s2$	$r2$ ($a1$)		$\begin{bmatrix} 1 & -\beta_1-1 & \beta_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\beta_2-1 & \beta_2 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \omega_{s1} \\ \omega_{a1} \\ \omega_{r1} \\ \omega_{s2} \\ \omega_{a2} \\ \omega_{r2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	$\frac{1+\beta_2}{1+\beta_1+\beta_2} \approx 3/5$	under drive
3*	$r1$ ($a2$)	$s1$	$r2$ ($a1$)		$\begin{bmatrix} 1 & -\beta_1-1 & \beta_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\beta_2-1 & \beta_2 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \omega_{s1} \\ \omega_{a1} \\ \omega_{r1} \\ \omega_{s2} \\ \omega_{a2} \\ \omega_{r2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	$\frac{\beta_1}{1+\beta_1} \approx 2/3$	under drive
4*†	$r1$ ($a2$)	$s2$	$r2$ ($a1$)		$\begin{bmatrix} 1 & -\beta_1-1 & \beta_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\beta_2-1 & \beta_2 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \omega_{s1} \\ \omega_{a1} \\ \omega_{r1} \\ \omega_{s2} \\ \omega_{a2} \\ \omega_{r2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	$\frac{1+\beta_2}{\beta_2} \approx 3/2$	over drive
5*	$s2$	$s1$	$r2$ ($a1$)		$\begin{bmatrix} 1 & -\beta_1-1 & \beta_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\beta_2-1 & \beta_2 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \omega_{s1} \\ \omega_{a1} \\ \omega_{r1} \\ \omega_{s2} \\ \omega_{a2} \\ \omega_{r2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	$\frac{\beta_1}{1+\beta_1+\beta_2} \approx 2/5$	under drive
6*†	$s2$	$r1$ ($a2$)	$r2$ ($a1$)		$\begin{bmatrix} 1 & -\beta_1-1 & \beta_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\beta_2-1 & \beta_2 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \omega_{s1} \\ \omega_{a1} \\ \omega_{r1} \\ \omega_{s2} \\ \omega_{a2} \\ \omega_{r2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	$-\frac{1}{\beta_2} \approx -1/2$	reverse under drive

States 3, 4, 5 and 6 (*) are respectively the high underdrive, overdrive, low underdrive and reverse states of the Axod 4R-speed automobile transmission. States 1, 2, 4 and 6 (†) are respectively the low underdrive, high underdrive, overdrive and reverse states of the HydraMatic 700 4R-speed automobile transmission.

two as their third and fourth equations because all these states are for the same hard link pair. Similarly, all states of Table 3 have the equations $\omega_{a1} - \omega_{r2} = 0$ and $\omega_{r1} - \omega_{a2} = 0$, due to their common hard link pair being $L1 = (a1, r2)$ and $L2 = (r1, a2)$.

The last two equations for a state correspond to its input and braked part in the usual way. For instance, since state 1 of Table 2 has braked and input parts $r1$ and $s1$ respectively, we add the corresponding equations $\omega_{r1} = 0$ and $\omega_{s1} = 1$. In this way, we obtain for this state the following complete set of six equations in six unknowns

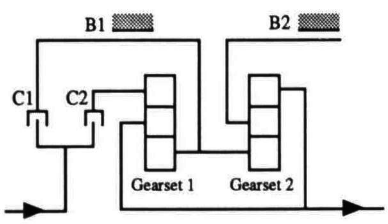
$$\begin{array}{lllll}
\omega_{s1} + (-\beta_1 - 1)\omega_{a1} + \beta_1\omega_{r1} & & & & = 0 \\
 & & \omega_{s2} + (-\beta_2 - 1)\omega_{a2} + \beta_2\omega_{r2} & & = 0 \\
\omega_{s1} & & -\omega_{s2} & & = 0 \\
 & \omega_{a1} & & -\omega_{r2} & = 0 \\
 & \omega_{r1} & & & = 0 \\
\omega_{s1} & & & & = 1.
\end{array}$$

As in Table 1, this set and its analogs for the other states appear in matrix form in Tables 2 and 3. It is these sets of simultaneous linear equations that are solved to obtain the expressions for the gear ratios ρ for the states. It is interesting to note that, in spite of using different ways of cross linking the two gearsets, the set of gear ratios of Table 3 are the same as those of Table 2.

2.2.7 *Two-planetary transmissions*

Section 2.2.5 discussed how transmissions could be built from a single planetary gearset with the addition of a switching network to allow incorporation of a range of the gearset's kinematic states. The latter single-gearset states were treated in Section 2.2.4. Analogously, the present section considers transmissions built from two linked planetary gearsets so as incorporate the two-planetary kinematic states of Section 2.2.6.

Figures 5 and 6 show how the two-planetary linking schemes of Tables 2 and 3 may be coupled with switching networks to obtain the well-known Simpson transmission (Ellinger, 1983; Husselbee, 1986) and



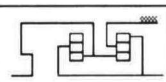
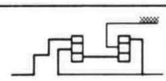
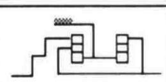
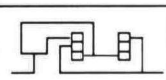
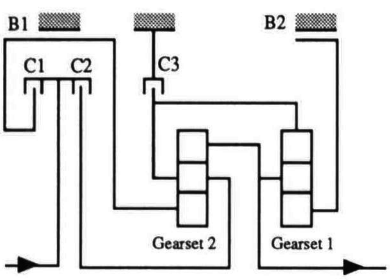
Speed	State	Clutches/Brakes C1	C2	B1	B2	Effective Transmission	In	Braked	Out	Ratio ρ	Ratio Type
0	2	On			On		$s1\,(s2)$	$a2$	$r2\,(a1)$	$-\frac{1}{\beta_2}$	reverse under drive
1	4		On		On		$r1$	$a2$	$r2\,(a1)$	$\frac{\beta_1}{1+\beta_1+\beta_2}$	under drive (low)
2	3		On	On			$r1$	$s1\,(s2)$	$r2\,(a1)$	$\frac{\beta_1}{1+\beta_1}$	under drive (high)
3	—	On	On				$r1\ \&\ s1\ (s2)$	—	$r2\,(a1)$	1	direct drive

FIGURE 5. A 3R-speed Simpson transmission based on two planetary gearsets. State numbers refer to Table 2



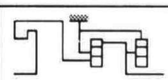
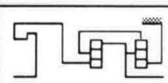
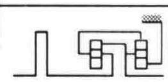
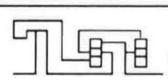
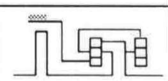
Speed	State	Clutches/Brakes C1	C2	C3	B1	B2	Effective Transmission	In	Braked	Out	Ratio ρ	Ratio Type
0	6	On		On				$s2$	$r1\,(a2)$	$r2\,(a1)$	$-\frac{1}{\beta_2}$	reverse under drive
1	5	On				On		$s2$	$s1$	$r2\,(a1)$	$\frac{\beta_1}{1+\beta_1+\beta_2}$	under drive (low)
2	3		On			On		$r1\,(a2)$	$s1$	$r2\,(a1)$	$\frac{\beta_1}{1+\beta_1}$	under drive (high)
3	—	On	On					$r1\,(a2)\ \&\ s2$	—	$r2\,(a1)$	1	direct drive
4	4		On		On			$r1\,(a2)$	$s2$	$r2\,(a1)$	$\frac{1+\beta_2}{\beta_2}$	over drive

FIGURE 6. A 4R-speed Axod transmission based on two planetary gearsets. State numbers refer to Table 3

https://doi.org/10.1017/S0890060400002651 Published online by Cambridge University Press

Axod transmission (Ford Motor Co., 1985) respectively. (Note the swap in placement of gearsets 1 and 2 in making the correspondence between the states of Table 3 and the Axod transmission of Figure 6. This is allowed because we have not made any real distinction between gearsets 1 and 2 other than in their names.) The Simpson transmission has three forward speeds and a reverse, while Axod has four forward speeds and a reverse. Unlike the Sturmey-Archer transmissions of Figures 3 and 4, the Simpson and Axod transmissions both achieve their set of gear ratios with a *fixed* output part $Out = r2$. As mentioned in Section 2.2.5, this was one of our motivations in going from a single-gearset transmission to the greater flexibility of a two-gearset transmission.

Note that the Axod transmission provides exactly the same ratios as the Simpson, plus an additional overdrive ratio $(1 + \beta_2)/\beta_2$. Since the Simpson linkage scheme of Table 2 also provides this overdrive ratio through state 5 there, one might expect that an 'extended Simpson' transmission could be built that adds this state 5 to those of the standard Simpson transmission in Figure 5. However, in practice such an extended Simpson transmission is not realizable because of topological constraints on the switching network that is needed. An extended Simpson transmission is in fact found by our algorithm, but only because we have not yet implemented the necessary topological constraints. The Axod transmission is also correctly rediscovered. (Both these will be seen in Figure 9 and Table 5 below.) We now look at how the automated discovery of these transmissions is made possible by formulation of the transmission design problem as a constraint satisfaction problem.

3. Variables and their domains

This section describes the variables and their domains of candidate values that we employ in formulating transmission design as a constraint satisfaction problem. The section after this will treat the corresponding constraints used. The domains and constraints given correspond to common practice in designing full-sized passenger vehicles in the United States.

Regarding *breadth* of our search space, remember we are restricting ourselves here to 4R-speed transmissions involving two planetary gearsets, cross linked by two hard links and no soft links. As for the Axod transmission of Figure 6, our transmissions will have one reverse speed, a low underdrive, a high underdrive, a direct drive and an overdrive. These speeds are denoted respectively by index values $i = 0$, 1, 2, 3 and 4. More generally, we use index value $i = d$ to denote direct drive and $i = n$ for the highest speed.

Regarding *depth* of our search space, for simplicity we currently ignore quite a few features of real transmissions. In particular, we do not consider the nature of the direct drive state nor the neutral state of our transmission designs. Nor will we be considering the topological or geometric nature of the paths linking gearsets between themselves and to the motor (for input), to the chassis (for braking) and to the wheels (for output), nor the nature of the clutches needed in these paths in order to switch between states (gear ratios). We will also not be considering the number of teeth on gears, which determines the actual numerical value of the gear ratios (as opposed to their symbolic algebraic form, which is the level at which these are considered here) and can effect whether a certain combination of gears can actually be assembled. Nor do we consider gear tooth shape, which can be important in reducing gear noise, gear wear and in improving power transfer efficiency. Most of these features however, we do expect to include in later formulations, as discussed in Section 7 below.

Leaving out the above-mentioned levels of detail in a transmission means that we are in effect designing abstractions or equivalence classes of transmissions rather than actual transmissions (although this is always true in some sense no matter how detailed one gets). It also means that even at our present level of abstraction, some of the solutions we obtain will probably be seen later to be invalid because they will not support any solutions at more detailed levels. In particular, we have still to study carefully the reasons that five of our seven solutions of Section 6 are probably unacceptable, but there is certainly no shortage of as yet unimplemented features (variables) and corresponding constraints that are likely to exclude them in a later more detailed formulation. For the CSP formulation of our present version of the transmission design problem, we will see that it suffices to have ten variables with an average domain size of about seven values each and a total of 43 constraints.

3.1 HARD LINKS (2 VARIABLES, 9 VALUES EACH)

A *link* is a connection between two gearset components that constrains those components to rotate at the same angular velocity. A *hard* (or *permanent* or *non-clutched*) link is a link built into the transmission so as to remain permanently in effect at all speeds. (Soft links are discussed below.) We assume here that the transmissions we are designing have exactly two planetary gearsets linked by exactly two hard links, L_1 and L_2.

For each of the hard links we need to decide what gearset components it joins together. The parts of gearset 1 that may constitute an end of a link are $Parts_1 = \{s1, a1, r1\}$, and the parts of gearset 2 are $Parts_2 = \{s2, a2, r2\}$. Note that planets are not candidates for the end of a link due to the difficulty of linking to these parts which, besides rotating on their axes, have centers which revolve about a sun gear. (See footnote 4, p. 143 and the related discussion.) A link is specified by the two parts which it links, and may thus be denoted as a pair $(p1, p2)$, with $p1 \in Parts_1$ and $p2 \in Parts_2$. In other words, the domains for the hard link variables $L1$ and $L2$ are

$$d_{L1} = d_{L2} = Parts_1 \times Parts_2$$
$$= \{(s1, s2)(s1, a2)(s1, r2)(a1, s2)(a1, a2)$$
$$\times (a1, r2)(r1, s2)(r1, a2)(r1, r2)\}. \quad (11)$$

The above domains assume that hard links between parts in a single gearset are not allowed. This is because such a link locks all gearset parts to rotate in synchrony, resulting in only a direct drive state ($\rho = 1$) being produced, as seen for example in state 8 of Table 1. This would make the pair of linked gearsets functionally no more flexible than a single gearset. Such a situation may be desirable *temporarily* for achieving one of the speeds of a transmission, as occurs for speed 2 of Figure 4 and for the third speeds of Figures 5 and 6. But we are here considering permanent links. There is no point in *permanently* linking parts on one gearset and thus effectively reducing a pair of gearsets permanently to one gearset.

3.2 SOFT LINKS (0 VARIABLES)

As opposed to hard (or permanent or non-clutched) links, *soft* (or *temporary* or *clutched*) links are not fixed for all speeds of a transmission, but may be in effect for some speeds and not in effect for others. The transition is achieved through clutches that re-configure connections between gearset parts. An example was seen in Figure 4, where a temporary link between the sun and arm could be established by engaging clutch C5 to give speed 2 and disengaging it to achieve speeds 1 and 3. Similarly the third speeds in Figures 5 and 6 are achieved by soft links created by simultaneously engaging clutches C1 and C2. Our current formulation of the transmission design problem assumes that there are no soft links between gearset parts.

3.3. OUTPUTS (0 VARIABLES)

In general, for each of the n forward speeds (in our case, $n = 4$) and the one reverse speed that we want our transmission to realize, we need to decide on the part to output from. This would introduce $n + 1$ new variables $Out_0, Out_1, Out_2, \ldots, Out_n$ into our CSP formulation. However, as explained in Section 2.2.5, it is common practice in automobile transmissions to have the same output part for each speed. Thus one output variable Out will suffice.

The domain of Out might *a priori* be thought to be $d_{Out} = Parts_1 \cup Parts_2 = \{s1, a1, r1, s2, a2, r2\}$. However, output from a sun is not used in practice due to stress considerations. The sun is the smallest diameter component of the relevant parts (sun, arm and ring). And since torque τ is the product of force F and the radius R about which the force is causing rotation, achieving a given torque requires a larger force if the radius is smaller (e.g. it is harder to shut a door by pushing near its hinges than by pushing at the door handle). Thus transmitting a given output torque from a sun gear would require a relatively large force, excessively stressing the sun gear teeth. Thus the domain of Out is reduced to $d_{Out} = \{a1, r1, a2, r2\}$.

But note that we are currently not distinguishing gearsets 1 and 2 other than by giving them different names. This will change in later formulations where one gearset will be considered closer to the motor, thus making the gearsets distinct at least with respect to the topological constraints which we will then be incorporating. But under the current formulation, outputing from gearset 1 is functionally equivalent to outputing from gearset 2. To avoid this kind of redundancy we arbitrarily choose gearset 2 as our default for output. The domain for Out thus becomes $d_{Out} = \{a2, r2\}$. And in fact, for simplicity in this initial formulation we have excluded $a2$ also, so that for the domain of Out we use simply $d_{Out} = \{r2\}$. The two-gearset states of Tables 2 and 3 and those of the transmissions in Figures 5 and 6 all used only this value for Out (whereas those of Table 1 and in the transmissions of Figures 3 and 4, allowed varying output for the one-gearset case shown there).

Note that since Out is here allowed only a single value, it is actually no longer a CSP variable as such, and need not be explicitly instantiated at some given level of the search tree. Thus our sample trace of

Figure 9 indicates no level for *Out*, which is globally taken as equal to $r2$. Analogous search trees for later formulations would have a level corresponding to instantiation of *Out* to at least $a2$ and $r2$.

3.4 INPUTS (4 VARIABLES, 6 VALUES EACH)

We are assuming that the kinematic state of our linked pair of planetaries at each speed is characterized, as described in Section 2.2.6, by the placement of two links, and the choice of input, output and braked parts. Unlike the links and the output part, the input and braked parts are allowed to vary with speed. To specify the input parts we therefore need a variable In_i for each speed $0 \leq i \leq n$. *A priori,* any of the combined six parts of the two gearsets is a potential input part at each speed, so that the domain of each variable In_i is

$$d_{In_i} = Parts_1 \cup Parts_2 = \{s1, a1, r1, s2, a2, r2\} \qquad \text{for } 0 \leq i \neq d \leq n.$$

We have excluded the $i = d$ case to make explicit that we are not presently concerned with the direct drive speed. Since $d = 3$ and $n = 4$ here, the above introduces four new variables In_0, In_1, In_2 and In_4 into our CSP formulation. These are the only In_i variables that appear for example in the trace of Figure 9.

3.5 BRAKES (4 VARIABLES, 6 VALUES EACH)

To specify the braked part at each speed we need a variable Br_i for $0 \leq i \leq n$. As with the inputs above, any of the combined six parts of the two gearsets is *a priori* a potential braked component at each speed so that the domain of each variable Br_i is

$$d_{Br_i} = Parts_1 \cup Parts_2 = \{s1, a1, r1, s2, a2, r2\} \qquad \text{for } 0 \leq i \neq d \leq n.$$

Again, as for inputs, we have excluded the $i = d$ case since we are not here concerned with the direct drive speed. For braked parts however the reason to exclude the $i = d$ case above is even stronger: there may not necessarily even be a braked part in direct drive. It may be replaced by an extra temporary link, as discussed for example in connection with state 8 of Table 1. Since $d = 3$ and $n = 4$ here, the above introduces four new variables Br_0, Br_1, Br_2 and Br_4 into our CSP formulation. These are the only Br_i variables that appear for example in the trace of Figure 9.

3.6 SIZES OF GEARS (0 VARIABLES)

In this initial formulation of the transmission design problem we do not concern ourselves directly with geometrical considerations such as the sizes of component gears. However gear size is relevant indirectly in that it effects the number of teeth allowed on the gears. In particular, it effects the ratio $\beta = T_r/T_s$, introduced in (2), between the number teeth on the ring gear and the number on the sun gear. Limits on gear sizes will hence place limits on β, and since all gear ratios can be expressed in terms of the β ratios of the transmission gearsets, gear size limits allow us to reason about the validity of gear ratios achieved. We will see examples of this below.

The following are the upper and lower bounds (in inches) that we will be assuming for the diameters of sun, planet and ring gears in a planetary gearset.

$$2 \leq D_s \leq 5 \qquad 1 \leq D_p \leq 2.5 \qquad 4 \leq D_r \leq 7. \qquad (12)$$

These arise as shown below, the first three cases being primary and the other three being deduced from them. We are assuming here a standard American passenger vehicle. The assumed and deduced values would be different for say a motorcycle or a truck. The basic logic would however be the same.

1. The ring gear cannot be greater than $D_r^{max} = 7$ in. This limit is imposed by the maximum space available in the vehicle for the transmission. (Having the ring gear fit of course ensures that the sun and planets fit since the ring encloses them.)
2. The sun gear cannot be smaller than $D_s^{min} = 2$ in. This is because in general the sun gear needs an input and/or output shaft running co-axially through its center, and 2 in. is the minimum sun diameter that allows such a shaft capable of handling the necessary loads and stresses.
3. A planet gear cannot be smaller than $D_p^{min} = 1$ in. This is so as to allow for a shaft (connecting the planet to the arm) to allow the planet to rotate on its axis. Note that the 1 in. diameter needed to allow for this is smaller that the 2 in. needed to allow for the sun gear's shaft, since the planet's shaft is just to allow rotation of the planet on its axis whereas the sun's shaft is also for input and/or output of torque and must therefore be larger.
4. The ring gear cannot be smaller than $D_r^{min} = 4$ in. This follows from $D_r = D_s + 2D_p$, by (3), and the

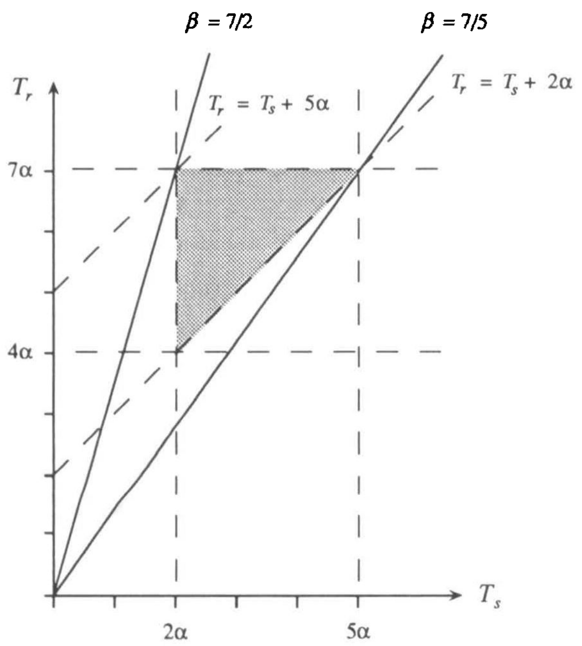


FIGURE 7. The region (gray) of possible pairs of teeth numbers T_r and T_s on the ring and sun gears of a planetary gearset. α denotes the dimetral pitch for the gearset. We see that the ratio $\beta = T_r/T_s$ must be in the range $7/5 \leq \beta \leq 7/2$.

above-mentioned minimum diameters of $D_s^{min} = 2$ and $D_p^{min} = 1$ for the sun and planet.

5. The sun gear cannot be greater than $D_s^{max} = 5$ in. This follows from $D_s = D_r - 2D_p$, by (3), and the above-mentioned maximum ring diameter $D_r^{max} = 7$ and minimum planet diameter $D_p^{min} = 1$.
6. The planet gear cannot be greater than $D_p^{max} = 2.5$ in. This follows from $D_p = (D_r - D_s)/2$, by (3), and the above-mentioned maximum ring diameter $D_r^{max} = 7$ and minimum sun diameter $D_s^{min} = 2$.

From these size limits and the use of (1) and (3), we can deduce, as shown graphically in Fig. 7, that β is bounded as follows

$$7/5 \leq \beta \leq 7/2. \tag{13}$$

4. Constraints

This section describes the constraints that we use in connection with the variables and values of the previous section in obtaining a CSP formulation of the transmission design problem. The division of the constraints into categories here is for the sake of providing some structure, but is not intended to be definitive. The constraints below are given labels of the form C_j or C_j'. These will be useful in our later sample trace in Figure 9, to identify which constraint is being applied where in the search process. Table 4 of Section 5 summarizes all our constraints and makes explicit their dependence on the CSP variables of Section 3. Remember that our current formulation does not consider the direct drive speed d in a transmission. In some cases below, we therefore explicitly exclude speed d when specifying a family of constraints. In other cases we do not, either for simplicity or because it is not yet clear if, or in what modified form, the constraints will apply for direct drive. In any case, speed-d versions of any constraints below are to be interpreted as not *necessarily* being valid, and of course none of them are used in our current search (as seen from Table 4 or Figure 9). The states in Tables 2 and 3 above, and in Table 5 below, all satisfy the applicable constraints of this section. So do the states of the exhaustive listing for all 18 hard link pairs given in Table A1 of the Appendix. They thus provide examples that may be useful in clarifying the meaning of the constraints below.

4.1 HARD LINK CONSTRAINTS

• *Non-connecting links*: As mentioned, we are interested in transmissions made of two planetary gearsets linked by two hard links (and no soft links). The allowed values for the two hard link variables $L1$ and $L2$ were given in (11). But not all pairs of hard link values are compatible. In particular, two hard links are not allowed to have a common end. If the two hard links are $L1 = (L1^-, L1^+)$ and $L2 = (L2^-, L2^+)$ then we can express this constraint as

$$L1^- \neq L2^- \quad \text{and} \quad L1^+ \neq L2^+. \tag{C$_1$}$$

Note that we do not have to explicitly enforce $L1^- \neq L2^+$ and $L1^+ \neq L2^-$; these cannot occur since $L1^-$ and $L2^-$ are both in gearset 1 and $L1^+$ and $L2^+$ are both in gearset 2, as discussed in Section 3.1. Actually, the reason given there for why the two ends of a hard link must be in different gearsets is closely related to the justification for the above constraint C_1 that the two hard links cannot have a common end. Consider for example two hard links $L1 = (L1^-, L1^+)$ and $L2 = (L2^-, L2^+)$ with a common end $L1^- = L2^-$ in gearset 1. Functionally this is like having a hard link between the other two non-common ends $L1^+$ and $L2^+$ in gearset 2, since these ends are constrained to move together due to their linkage via the common part in gearset 1. We have then, indirectly, what

https://doi.org/10.1017/S0890060400002651 Published online by Cambridge University Press

amounts to a hard link within a single gearset—and we have seen in Section 3.1 that such a situation is undesirable, as it permanently reduces a two-gearset configuration to essentially a one-gearset configuration.

• *Link renaming equivalence*: In linking gearsets using two links it does not matter which link we call $L1$ and which we call $L2$. That is, having two links $L1=(L1^-, L1^+)$ and $L2=(L2^-, L2^+)$ is physically the same as having the two links $L1=(L2^-, L2^+)$ and $L2=(L1^-, L1^+)$. This kind of redundancy can be avoided by requiring say link $L1$ to be lexographically less than link $L2$, with respect to the underlying order of parts given in (9). We write this constraint as

$$(L1^-, L1^+) < (L2^-, L2^+). \qquad (C_2)$$

Using constraints C_1 and C_2 reduces the *a priori* $81 = 9\times 9$ possible pairs of links for two planetaries down to the 18 pairs shown in Figure 8.

• *Link end equivalence*: Since two parts joined by a link are by definition constrained to move together, braking one end of a link also brakes the other end. Similarly, inputing to, or outputing from, one end of a link is respectively like inputing to, or outputing from, the other end. This kind of functional redundancy can be avoided by choosing say the gearset 1 end of a link as the preferred part, and not allowing the gearset 2 end as legal for the input part In_i or braked part Br_i at any speed i. If the two hard links are $L1=(L1^-, L1^+)$ and $L2=(L2^-, L2^+)$, then we can express these constraints as

$$In_i \neq L1^+ \quad \text{for } 0\le i\le n. \qquad (C^i_{3A})$$

$$In_i \neq L2^+ \quad \text{for } 0\le i\le n. \qquad (C^i_{3B})$$

$$Br_i \neq L1^+ \quad \text{for } 0\le i\le n. \qquad (C^i_{4A})$$

$$Br_i \neq L2^+ \quad \text{for } 0\le i\le n. \qquad (C^i_{4B})$$

Note that the above constraints remove functionally equivalent states. Those states however will not necessarily remain equivalent in other ways in later formulations. In particular, when we distinguish gearsets 1 and 2 by their placement with respect to the motor the two ends of a link will generally become topologically inequivalent. However at the current level of formulation, we are justified in assuming equivalence of ends and using the above constraints. Note that we do not need constraints analogous to those above for restricting *Out* to a preferred end of each link, because *Out* is assumed fixed here at $Out = r2$ and is thus already restricted to the gearset 2 end of any link it impinges on.

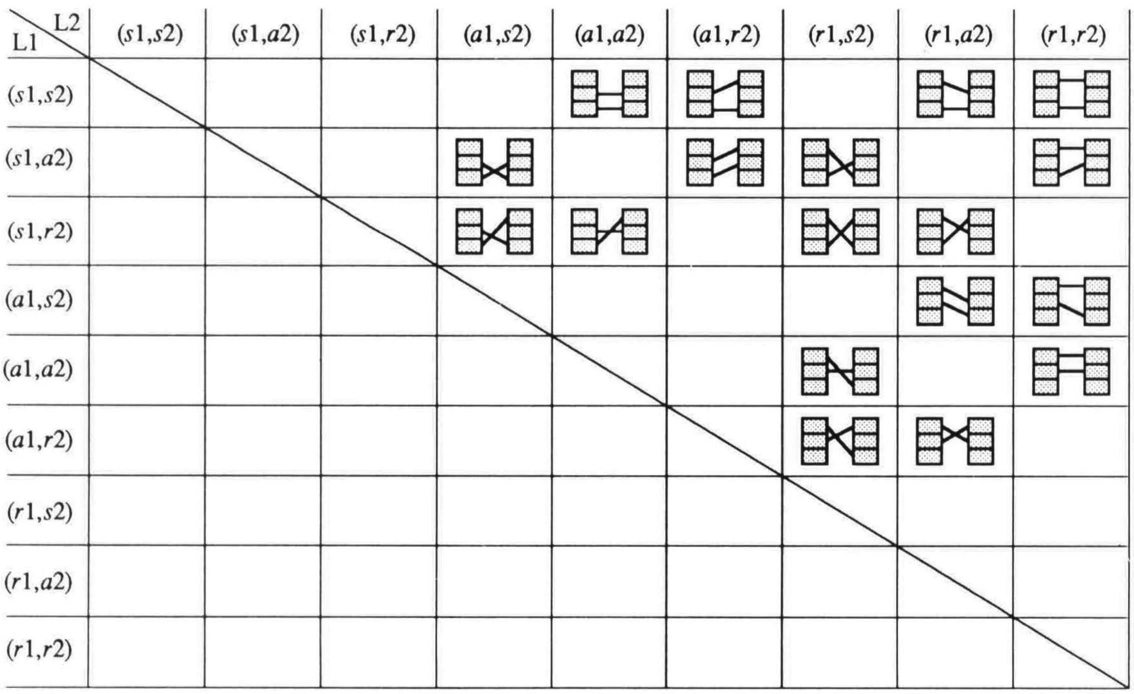


FIGURE 8. The 18 legal (out of $9\times 9 = 81$ possible) pairs of hard links between two planetary gearsets.

https://doi.org/10.1017/S0890060400002651 Published online by Cambridge University Press

4.2 GEAR RATIO OR SPEED CONSTRAINTS

• *Don't brake the output*: It is obvious that a brake should not be applied to the output part, else we get no output torque. Thus we have the constraints

$$Br_i \neq Out \qquad \text{for } 0 \leq i \leq n. \qquad (C_5^i)$$

Since braking a part that is linked to the output causes the same problem—no output torque—we interpret C_5^i to mean that neither the output part, nor a part joined to it by a link, may be braked.

• *Don't brake the input*: As with not braking the output part above, we also cannot brake the input part, else we get no input torque. Thus we have the constraints

$$Br_i \neq In_i \qquad \text{for } 0 \leq i \leq n. \qquad (C_6^i)$$

Unlike with constraints C_5^i, we need not here be concerned with also not braking a part linked to the input part. Such a situation cannot arise since as long as constraints C_{3A}^i, C_{3B}^i, C_{4A}^i and C_{4B}^i are checked before C_6^i constraints. The former constraints ensure that if either the input part or the braked part is on a link then they are at the gearset 1 end of that link, and cannot thus be on opposite ends of a link.

• *Don't input to the output*: When the input part is the same as the output part, the gear ratio $\rho = \omega_{Out}/\omega_{In}$ must of course be $\rho = 1$ (see for example state 7 of Table 1), so we get a direct drive speed. Thus for all non-direct-drive speeds this possibility must be excluded. We therefore have the constraints

$$In_i \neq Out \qquad \text{for } 0 \leq i \neq d \leq n \qquad (C_7^i)$$

where as before, index d is the index of the direct drive speed (for which case it *is* allowable to have $In_d = Out$). Since inputing to a part that is linked to the output causes the same problem—a direct drive ratio—we interpret C_7^i to mean that neither the output part, nor a part joined to it by a link, may be the input part.

The above constraints C_1 to C_7^i are the formal reason why, as seen in Tables 2, 3 and A1, there are six functionally different non-direct-drive states for each choice of output part and hard link pair. The functional equivalence of the two ends of a link, as implicit in C_{3A}^i, C_{3B}^i, C_{4A}^i and C_{4B}^i, reduce the two end parts of a link in effect to a single 'part'. Since there are two hard links, with no ends in common by C_1, the six parts of the two gearsets are thus functionally equivalent to only $6 - 2 \times 1 = 4$ parts. When one of these is chosen as the output part *Out*, there are only three choices left for the input part *In* and braked part *Br* because neither *In* nor *Br* may equal *Out* by C_5^i and C_7^i. Thus there are three choices for *In* and for each of these, since *Br* may not equal *In* by C_6^i, there are a remaining two choices for *Br*. We thus have a total of $3 \times 2 = 6$ choices for the *In*/*Br* configuration for a specified output part. Of course, since there are in general four functionally inequivalent choices in the first place for *Out* (although here we use only $Out = r2$), there are thus $4 \times 3 \times 2 = 24$ possible *In*/*Br*/*Out* configurations for each hard link combination of two gearsets, as opposed to only the six of Table 1 for a single gearset.

• *Different ratios in different gears*: By definition, different speeds or 'gears' must have different gear ratios. Since we are assuming a fixed output part *Out* in each speed i, the gear ratio in speed i (for a given pair of hard links) varies only with the input part In_i and the braked part Br_i. To avoid the same ratio at different speeds i and j, we thus require at least one of these variables to be different, giving the constraints

$$(In_i \neq In_j) \text{ or } (Br_i \neq Br_j) \qquad \text{for } 0 \leq i < j \leq n. \qquad (C_8^{i,j})$$

• *Gear ratio ranges*: In designing a transmission, perhaps the most basic features are the number of speeds and the values of gear ratio available at those speeds. The number of speeds desirable in a transmission depends on various factors such as the type of vehicle, where and how it will be used, the power of the motor and how important it is to save fuel and reduce engine wear. A nice discussion of these considerations appears in Heldt (1955), pp. 5 and 6. We are assuming here a transmission with 1 reverse speed ($\rho_0 < 0$) and $n = 4$ forward speeds consisting of two underdrives ($0 < \rho_1 < \rho_2 < 1$), a direct drive ($\rho_3 = 1$) and an overdrive ($\rho_4 > 1$) speed. An example of such a 4R-speed transmission, Axod, was seen in Figure 6. Analogous transmissions with fewer speeds were shown in Figures 3 to 5.

The following constraints specify the ranges we will consider acceptable for the gear ratios ρ_i at these speeds. They represent common practice for a 4R-speed automatic transmission in a full-sized passenger vehicle. Note that the ratio for direct drive is included only for completeness. As mentioned, we do not explicitly concern ourselves with the direct drive state in our present formulation.

Reverse Gear Ratio:

$$-\tfrac{2}{5} \leq \rho_0(\beta_1, \beta_2; L1, L2, In_0, Br_0) \leq -\tfrac{1}{5} \qquad (C_9)$$

https://doi.org/10.1017/S0890060400002651 Published online by Cambridge University Press

First Gear Ratio (low underdrive):

$$\tfrac{1}{3} \leq \rho_1(\beta_1, \beta_2; L1, L2, In_1, Br_1) \leq \tfrac{1}{2} \qquad (C_{10})$$

Second Gear Ratio (high underdrive):

$$\tfrac{3}{5} \leq \rho_2(\beta_1, \beta_2; L1, L2, In_2, Br_2) \leq \tfrac{4}{5} \qquad (C_{11})$$

Third Gear Ratio (direct drive):

$$\rho_3 = 1 \qquad (C_{12})$$

Fourth Gear Ratio (overdrive):

$$1\tfrac{1}{4} \leq \rho_4(\beta_1, \beta_2; L1, L2, In_4, Br_4) \leq 1\tfrac{2}{3} \qquad (C_{13})$$

For each kinematic state that our algorithm generates, it computes $\rho_i(\beta_1, \beta_2)$ symbolically as a function of β_1 and β_2, the ring-to-sun teeth number ratios for gearsets 1 and 2. These symbolic expressions are obtained by solving a corresponding set of simultaneous linear (kinematic) equations, as discussed in connection with Tables 2 and 3. The form of these equations depends on the particular hard links used (the values of $L1$ and $L2$) and on the parts being used for input and for grounding at speed i (the values of In_i and Br_i). The ratio ρ_i obtained from these equations—but not the form of the equations *per se*—depends further on the choice of output part at speed i (the value of Out_i). Thus $\rho_i(\beta_1, \beta_2)$ is more fully denoted as $\rho_i(\beta_1, \beta_2; L1, L2, In_i, Br_i, Out_i)$ or, in our case, $\rho_i(\beta_1, \beta_2; L1, L2, In_i, Br_i)$ since we consider the output part fixed at $Out_i = r2$ for all speeds. It is in this way that (C_9) to (C_{13}) formally act as CSP constraints on the actual variables of our CSP formulation. This is made explicit in the corresponding rows of Table 4 below. Examples of obtained symbolic expressions $\rho_i(\beta_1, \beta_2; L1, L2, In_i, Br_i)$ were seen in Tables 2 and 3 and in Figures 5 and 6. Others are seen below in Table 5 and in the Appendix.

The above ratio-range constraints C_9 to C_{13} are thus somewhat different than the other constraints of our formulation, in that they involve our CSP variables only implicitly, but are explicit functions of auxiliary variables β_1 and β_2. Certainly the gear ratios ρ_i are not themselves CSP variables, and the corresponding constraints above cannot be tested simply by using instantiated values of ρ_i. A ρ_i expression is a function of β_1 and β_2, which in turn are functions of the teeth numbers T_{s1}, T_{s2}, T_{r1} and T_{r2}. So values for ρ_i could be obtained by instantiating these as CSP variables; their domains would be discrete and finite as required. This would be a natural approach, but as discussed in Section 3, in our current formulation we do not wish to go to this level of detail and corresponding search complexity. So teeth numbers for gears are not features (CSP variables) of our design at present, and the above ratio-range constraints must be tested less directly.[7]

The indirect test we use is based on the work of Ward (1988). Given a function $z = f(x, y)$, monotonic in both x and y, it is possible to bound the variation in z given bounds on the variation in x and y. In particular, if $a \leq x \leq b$ and $c \leq y \leq d$, then

$$\min\{f(a, c), f(a, d), f(b, c), f(b, d)\} \leq z \leq \max\{f(a, c), f(a, d), f(b, c), f(b, d)\}. \qquad (14)$$

Given that we have bounds on β_1 and β_2 from (13), we can use (14) to bound the value of $\rho(\beta_1, \beta_2)$ for a given kinematic state.[8] For example, state 6 of Table 2 results in a gear ratio of $\rho = (1 + \beta_2)/(1 + \beta_1 + \beta_2)$, which is monotonic in both β_1 and β_2. We thus have, by (14) and (13), that

$$\min\{0.632, 0.407, 0.763, 0.563\} \leq \rho(\beta_1, \beta_2) \leq \max\{0.632, 0.407, 0.763, 0.563\}$$

or

$$0.407 \leq \rho(\beta_1, \beta_2) \leq 0.763$$

We can use this range for ρ to conclude that constraints C_9, C_{12} and C_{13} cannot be satisfied, since there is no overlap of the range with the ranges required by those constraints. Thus the corresponding state cannot be used to provide a reverse, direct drive or overdrive gear ratio. Constraints C_{10} and C_{11} however are not violated, since there is overlap of the range $[0.407, 0.763]$ with that required by each of those constraints, and the state may still be used to provide a low or high underdrive ratio. (It is this kind

[7] Even if we did consider teeth numbers, it would be preferable in terms of combinatorial explosion to postpone instantiation of the corresponding variables till the last levels of the search tree—and an initial application of our indirect scheme for testing the ratio-range constraints would still be desirable to allow some earlier pruning higher up in the tree.

[8] This assumes $\rho(\beta_1, \beta_2)$ is monotonic in β_1 and β_2 as required by (14). We have found that of the $18 \times 6 = 108$ states that our search generates, all but four do have monotonic ρ functions. (See the discussion in the Appendix.) We hope to generalize (14) appropriately in the future to handle the remaining non-monotonic cases. Even if the generalization cannot be made it is not critical, since once we refine the formulation to include teeth number, as we intend to do, we may test constraints C_9 to C_{13} numerically as mentioned above. In that case the approach of (14) would still be useful, but only because it allows us to partially test the ratio-range constraints earlier for a more efficient search, as implied in footnote 7 above.

of numerical range information, and corresponding allowable speeds, that is given exhaustively in the appendix. The above example for instance, appears there as state 6 of case 2.)

Note that non-overlap of a state's range for ρ and that required by a ratio-range constraint, means that the constraint is definitely violated. Overlap on the other hand, does not mean that the constraint can be necessarily satisfied. It just means that the ratio-range constraint is not definitely violated. The ratio-range constraint is only definitely satisfied when a specific value of ρ is found in the range required by the constraint. Such a full check of satisfaction is not possible under the current formulation because we do not get to the level of assigning gear ratios specific numerical values, since we do not consider teeth numbers *per se*. In this sense the use of Ward's method in the current implementation is a way of *partially* checking the ratio range constraints.

- *Gear ratio span and step*: Constraints C_9 to C_{13} are constraints on individual ratios (and hence on the corresponding $L1$, $L2$, In_i and Br_i variables). But the value of one ratio may effect the allowable values for other ratios. Two important classes of inter-ratio constraints are the *span* constraints and the *step* constraints. The former specify what the spread between the highest and lowest gear ratio should be. The latter specify what the ratio between successive gear ratios should be. We leave implementation of these constraints to a later stage. They will be a lot easier to incorporate after teeth numbers have been decided (see Section 4.4) so that corresponding gear ratios will be available numerically, rather than just symbolically as here. Span and step issues are discussed in (Heldt, 1955), pages 6 to 11.

4.3 SIMPLICITY-OF-SWITCHING CONSTRAINTS

An important class of constraints are what we call the simplicity-of-switching constraints. This notion was introduced in Section 2.2.5. The idea is to keep to a minimum the necessary clutch and brake switches in changing the input and/or braked parts between 'directly interchangeable' pairs of speeds, i.e. between pairs of speeds for which a direct move from one to the other is likely to occur in practice. If the clutch and brake pattern is required to change too much between such speeds, it becomes very difficult to build the switching network so as to effect these changes in synchrony. In that case one risks the potentially dangerous possibility that unintended speeds will be entered transiently before a given target speed is achieved.

Thus we allow only simple changes of braking and clutching between directly interchangeable speeds. In particular, we allow at most one of the braked part and the input part to differ in the states for such a pair of speeds. This gives us the following simplicity-of-switching constraints

$$(In_i = In_j) \text{ or } (Br_i = Br_j)$$
$$\text{for } (i, j) \in \{(0, 1), (1, 2), (2, 4)\}. \qquad (C_{14}^{i,j})$$

The most basic pairs of directly interchangeable speeds are the pairs of *successive* speeds: (1, 2), (2, 3) and (3, 4). The latter two pairs are excluded above simply because our transmission designs currently ignore speed 3 (direct drive). Actually, in (Nadel and Lin, 1991*a*) and the IJCAI-91 workshop version of (Nadel and Lin, 1991*b*), we only used the above $(i, j) = (1, 2)$ case, and did not use the $(i, j) = (0, 1)$ and $(i, j) = (2, 4)$ cases. This was due to a misunderstanding over what is common practice at Ford. However, it seems that it is not uncommon for drivers to need direct switches between speeds (0, 1) and also between speeds (2, 4). For example, when a vehicle is stuck in snow or mud, it is often useful to rock it loose by repetitively moving directly between reverse ($i = 0$) and first speed ($j = 1$) and back again, even though usually one would stop off at neutral in between. The present use of the additional $(i, j) =$ (0, 1) and $(i, j) = (2, 4)$ cases is the reason that (as we will see) there are only seven solutions found here while in (Nadel and Lin, 1991*a*) and the IJCAI-91 version of (Nadel and Lin, 1991*b*) 10 solutions were found (see footnote 2, p. 138). It is in fact a simple matter to check visually that three of those earlier 10 solutions, numbers 2, 3 and 7, violated either the $(i, j) = (0, 1)$ or the $(i, j) = (2, 4)$ case of constraints $C_{14}^{i,j}$ and hence should be excluded as solutions for our present version of the transmission design problem.

4.4 GEARING CONSTRAINTS

Gearing constraints are those which directly effect any aspect of the gears making up the planetary gearsets in our transmissions. This includes gear tooth shape, gear weight, size, etc. However, at least for quite a while, we will not be designing transmissions at that level of detail. At most we will be interested in the number of teeth on the various gears, because of their effect on gear ratios ρ_i via parameters $\beta_1 = T_{r1}/T_{s1}$ and $\beta_2 = T_{r2}/T_{s2}$. Even just for teeth numbers on gears there are quite a few constraints that apply. Some are 'hard' constraints which must be satisfied, such as those which ensure that the component gears of an assembled gearset can mesh with each other (Kelly, 1959). Others are 'soft' constraints which it is desirable, but not essential, to satisfy such as those which reduce gear tooth noise (Jones and Route, 1963; Route, 1988) or gear tooth wear (Jones, 1988). In any case, at the present level in our design of transmissions we do not consider teeth number or other gear-level features, and postpone a description of gearing constraints till a later paper (Nadel, 1991*b*).

4.5 TOPOLOGICAL CONSTRAINTS

Another class of constraints that becomes relevant when we refine our design of transmissions to the level seen in Figures 3 to 6 are the *topological constraints*. These are necessary in order to ensure that the proposed hardware paths linking gears and clutches and brakes are topologically acceptable. Loosely speaking, this means that the paths do not pass through each other, or require other contortions which are not acceptable. As with gearing constraints, since in the work reported here we are not refining our designs to this level, we postpone a description of topological constraints till a later paper (Nadel, 1991*c*).

5. Some theoretical considerations

The previous two sections gave our CSP formulation for the version of the transmission design problem we are currently considering. This section summarizes our formulation and looks at it in relation to some theoretical issues treated in earlier work. The following table summarizes the CSP variables we use, as introduced in Section 3.

Variable z_i	$L1$	$L2$	In_0	Br_0	In_1	Br_1	In_2	Br_2	In_4	Br_4
Domain size m_{z_i}	9	9	6	6	6	6	6	6	6	6

We see there are 10 variables, two with domain sizes of 9 and eight with domain sizes of 6. So the overall Cartesian product $D = \times_{z_i \in Z} d_{z_i}$, or search space, from which solutions are to be found has size

$$M = |D| = |\times_{z_i \in Z} d_{z_i}| = \prod_{z_i \in Z} |d_{z_i}| = \prod_{z_i \in Z} m_{z_i}$$
$$= 9^2 \times 6^8 = 136{,}048{,}896.$$

Table 4 summarizes the constraints we use, as introduced in Section 4. For each constraint C_j, the table shows by a × the variables z_i making up its argument set Z_j. Also shown for each constraint is:

- its arity $A_j = |Z_j|$ (number of argument variables);
- the size $M_j = |D_j| = \prod_{z_i \in Z_j} m_{z_i}$ of its Cartesian product $D_j = \times_{z_i \in Z_j} d_{z_i}$;
- its satisfiability S_j (the number of tuples from D_j that satisfy C_j);
- its satisfiability ratio $R_j = S_j/M_j$ (the fraction of tuples from D_j that satisfy C_j).

The above notation was introduced in Section 2.1. We see from Table 4 that there are $c = 43$ constraints, of which eight are unary ($A_j = 1$), 22 are binary ($A_j = 2$) and 13 are quaternary ($A_j = 4$). With $n = 10$ variables, of average domain size $6.6 = (2 \times 9 + 8 \times 6)/10$, and with $c = 43$ mostly binary constraints, we see that our formulation has resulted in a CSP instance that is in fact surprisingly small. It is of about the same order of difficulty as the 10-queens problem (of finding all ways to place 10 queens on a 10×10 chessboard so that no two queens attack each other) under the standard CSP formulation (Nadel, 1990*a*), which has ten variables, each of domain size 10, subject to $\binom{10}{2} = 45$ binary constraints.

The 10-queens problem has 724 solutions (Nadel, 1989). How many can we expect our present problem to have, from its underlying search space of over 136 million candidate solutions? It has been shown (Nadel, 1986, 1991*a*) that the expected number of solutions for an instance in what we call a 'small class' of CSP instances, is given by

$$\bar{S} = \left[\prod_{z_i \in Z} m_{z_i}\right]\left[\prod_{j=1}^{c} R_j\right]. \tag{15}$$

This is a slight generalization of equation (36) in (Nudel, 1983*a*). It considers all instances of the underlying small class to be equally likely. A small class is the set of all CSP instances having a given set Z of n variables z_i, each of corresponding domain size m_{z_i} and having c constraints C_j each of corresponding satisfiability ratio $R_j = S_j/M_j$. There are in fact $\prod_{j=1}^{c} \binom{M_j}{S_j}$ problem instances in such a class, because for each of the c constraints there are $\binom{M_j}{S_j}$ ways to choose the required S_j satisfying tuples from the M_j tuples of the Cartesian product D_j for the jth constraint C_j. Clearly, in spite of the name 'small class', such classes can be extremely large. In our case, using the M_j and S_j values from Table 4, this gives

$$\binom{81}{36}^2 \cdot \binom{54}{45}^{16} \cdot \binom{6}{5}^8 \cdot \binom{36}{30}^4 \cdot \binom{1296}{1260}^6 \cdot \binom{2916}{918} \cdot \binom{2916}{378}^2 \cdot \binom{2916}{864} \cdot \binom{1296}{396}^3 \approx 4.6 \times 10^{4218}$$

for the size of the small class to which our problem belongs. Using (15) with the domain sizes m_{z_i} from the

https://doi.org/10.1017/S0890060400002651 Published online by Cambridge University Press

TABLE 4. The 43 constraints C_j of our CSP formulation, the variables (denoted by $\times$) of their argument sets Z_j, their arities A_j, Cartesian product sizes M_j, satisfiabilities S_j and satisfiability ratios R_j

j	Constraint C_j	variables $L1$	$L2$	In_0	Br_0	In_1	Br_1	In_2	Br_2	In_4	Br_4	A_j	M_j	S_j	$R_j = S_j/M_j$
1	C_1	×	×									2	81	36	4/9
2	C_2	×	×									2	81	36	4/9
3	C^0_{3A}	×		×								2	54	45	5/6
4	C^1_{3A}	×				×						2	54	45	5/6
5	C^2_{3A}	×						×				2	54	45	5/6
6	C^4_{3A}	×								×		2	54	45	5/6
7	C^0_{3B}		×	×								2	54	45	5/6
8	C^1_{3B}		×			×						2	54	45	5/6
9	C^2_{3B}		×					×				2	54	45	5/6
10	C^4_{3B}		×							×		2	54	45	5/6
11	C^0_{4A}	×			×							2	54	45	5/6
12	C^1_{4A}	×					×					2	54	45	5/6
13	C^2_{4A}	×							×			2	54	45	5/6
14	C^4_{4A}	×									×	2	54	45	5/6
15	C^0_{4B}		×		×							2	54	45	5/6
16	C^1_{4B}		×				×					2	54	45	5/6
17	C^2_{4B}		×						×			2	54	45	5/6
18	C^4_{4B}		×								×	2	54	45	5/6
19	C^0_5				×							1	6	5	5/6
20	C^1_5						×					1	6	5	5/6
21	C^2_5								×			1	6	5	5/6
22	C^4_5										×	1	6	5	5/6
23	C^0_6			×	×							2	36	30	5/6
24	C^1_6					×	×					2	36	30	5/6
25	C^2_6							×	×			2	36	30	5/6
26	C^4_6									×	×	2	36	30	5/6
27	C^0_7			×								1	6	5	5/6
28	C^1_7					×						1	6	5	5/6
29	C^2_7							×				1	6	5	5/6
30	C^4_7									×		1	6	5	5/6
31	$C^{0,1}_8$			×	×	×	×					4	1296	1260	35/36
32	$C^{0,2}_8$			×	×			×	×			4	1296	1260	35/36
33	$C^{1,2}_8$					×	×	×	×			4	1296	1260	35/36
34	$C^{0,4}_8$			×	×					×	×	4	1296	1260	35/36
35	$C^{1,4}_8$					×	×			×	×	4	1296	1260	35/36
36	$C^{2,4}_8$							×	×	×	×	4	1296	1260	35/36
37	C_9	×	×	×	×							4	2916	≈ 918	$\approx 34/108$
38	C_{10}	×	×			×	×					4	2916	≈ 378	$\approx 14/108$
39	C_{11}	×	×					×	×			4	2916	≈ 378	$\approx 14/108$
40	C_{13}	×	×							×	×	4	2916	≈ 891	$\approx 33/108$
41	$C^{0,1}_{14}$			×	×	×	×					4	1296	396	11 / 36
42	$C^{1,2}_{14}$					×	×	×	×			4	1296	396	11 / 36
43	$C^{2,4}_{14}$							×	×	×	×	4	1296	396	11 / 36

table above and the satisfiability ratios R_j from Table 4, the expected number of solutions for an instance in this small case is found to be

$$\bar{S} = 9^2 \cdot 6^8 \cdot (\tfrac{4}{9})^2 \cdot (\tfrac{5}{6})^{28} \cdot (\tfrac{35}{36})^6 \cdot \tfrac{34}{108} \cdot (\tfrac{14}{108})^2 \cdot \tfrac{33}{108} \cdot (\tfrac{11}{36})^3$$
$$= 6.35.$$

In the next section we will see that our problem in fact has seven solutions. The theoretical expected number of solutions, 6.35, for the parent small class is remarkably close to the actual number, 7, for our problem given that its search space has over 136 million candidate solutions, and given that by definition an expected number is an average over a whole class (which in our case contains about 4.6×10^{4218} instances) and is not meant to be exact for individual subsumed instances.

However, if the instances of a class over which an expectation is obtained are reasonably similar, then a class average of a quantity may be expected to be a good approximation to the values of that quantity for individual subsumed instances. It is in fact for this reason that we have chosen to analyse CSP in terms of small classes: such classes contain instances similar enough that their *time complexities* are reasonably similar and thus the small class average complexity provides a useful approximation of the complexity of solving most individual instances of that small class. This has been found empirically (Nadel, 1986, 1991*a*; Nudel, 1983*a*) to be so for the complexity of solving CSP instances by the Backtracking algorithm and by the Forward Checking algorithm. Experiments have shown that for both these algorithm, about 80% of the instances in a small class have time complexity within 15% of their small class expected time complexity. The number of solutions (but not necessarily the time complexity) of our example is for instance $10.2\% = (7 - 6.35)/6.35$ from the expected value for its small class. A problem (sub)class with this kind of similarity for all, or most, of its member instances we call 'homogeneous' with respect to the quantity of interest. It is a useful goal to aim for in the analysis (best-case, worst-case or expected-case) of any random variable. It is achieved by making the appropriate ('homogenizing') choice of parameters in terms of which the sample space is to be analysed. A good choice of analysis parameters partitions the outcomes into subclasses that are each homogeneous in the quantity of interest. A consideration of homogeneity in the context of analysing the well-known insertion sort algorithm appears in (Nadel, 1988). CSP small classes are less homogeneous with respect to the quantity *number of solutions* than they are with respect to Backtracking and Forward Checking *time complexity*. Nevertheless, even for number of solutions, the average value for the parent small class is often a useful predictor of the number of solutions for an individual instance, as seen in our present example.

How were the parameter values in Table 4 found? The A_j and M_j values are of course readily obtained from their definitions. The constraint satisfiability values S_j are less easy to obtain generally. By definition S_j is the number of tuples that satisfy the constraint C_j from the total number M_j in the Cartesian product D_j of that constraint. As such they can always be obtained by considering each member of D_j (which is finite of course, since all domain sizes are finite) and testing it with respect to C_j. This counting process can be tedious, but may often be avoided or at least streamlined by a little mathematical analysis. Due to space limitations, and since it is relatively tangential to the theme of this paper, we forego discussion of the techniques applicable for our present constraints. In any case, given S_j and M_j values, the constraint satisfiability ratios in Table 4 are then found as simply $R_j = S_j/M_j$.

The S_j and R_j values for four constraints, C_9, C_{10}, C_{11} and C_{13} (actually constraints C_{37} to C_{40} according to the index j in the left column of Table 4), do however require some further discussion. Each of these constraints is in terms of the gear ratio ρ for a state where, as described, a state's gear ratio is obtained by solving a set of simultaneous equations characterizing that state. A state corresponds to a 4-tuple ($L1$, $L2$, In, Br), as discussed in Section 4.2 in connection with the ratio-range constraints. There are here $2916 = 9 \times 9 \times 6 \times 6$ such combinations possible *a priori*. Counting S_9, S_{10}, S_{11} and S_{13}, from first principles would therefore require solving a set of simultaneous equations for each of these 2916 combinations. This is an unappealing approach, and no mathematical short cut is obvious. What we have done therefore is to use only the $108 = 18 \times 6$ combinations of ($L1$, $L2$, In, Br) listed in Table A1 of the Appendix, for each of which we *have* solved the corresponding set of simultaneous equations to obtain its gear ratio symbolically and have tested this gear ratio with respect to each of the ratio-range constraints C_9, C_{10}, C_{11} and C_{13} to determine which speeds the state can provide. We consider the 108 combinations of Table A1 as representative of the full 2916 combinations in the sense that both the full and reduced subset have an approximately equal fraction that satisfy constraint C_9, and similarly for C_{10}, C_{11} and C_{13}. This is not an implausible assumption, even though it is true that the 108 combinations are 'biased' in that they are the particular ones which satisfy constraints C_1 to C_7^i; see the related discussion in Section 4.2 after constraint C_7^i. This bias does not necessarily effect the *proportion* of states which satisfy constraints C_9, C_{10}, C_{11} and C_{13} in the subset compared to the full set. As shown in Table A1, of the 108 states appearing, 34, 14, 14 and 33 are valid for speeds 0, 1, 2 and 4 respectively and hence satisfy constraints C_9, C_{10}, C_{11} and C_{13} respectively. We assume the full set of 2916 has similar proportions of satisfying states, so that $R_9 \approx 34/108$, $R_{10} \approx 14/108$, $R_{11} \approx 14/108$ and $R_{13} \approx 33/108$. For completeness of Table 4, we translate these back to the corresponding approximate S_j values for the full set as follows: $S_9 \approx 918 = 34/108 \times 2916$, $S_{10} \approx 378 = 14/108 \times 2916$, $S_{11} \approx 378 = 14/108 \times 2916$ and $S_{13} \approx 891 = 33/108 \times 2916$.

6. Our results

Sections 3 and 4 showed how we can formulate the transmission design problem as a constraint satisfaction problem, specifying respectively the CSP variables and domains which we use, and the CSP constraints. In this section we see how the resulting

https://doi.org/10.1017/S0890060400002651 Published online by Cambridge University Press

CSP instance is solved and discuss the solutions which are found.

6.1 THE SEARCH TREE

There are many algorithms for solving constraint satisfaction problems, as seen for example in (Dechter and Pearl, 1988; Haralick and Elliot, 1980; Mackworth, 1987; Nadel, 1989). In the initial stages of our work we used the *Forward Checking* algorithm because experiments have shown it to be one of the most efficient CSP algorithms (Haralick and Elliot, 1980; Nadel, 1989). However, as discussed in Sections 2.1 and 5, our current formulation of the transmission design problem results in a rather simple CSP instance, about as difficult as the 10-queens problem. As such, the extra efficiency afforded by algorithms such as Forward Checking is not particularly important yet, and the easier-to-implement Backtracking algorithm (Haralick and Elliot, 1980), (Nadel, 1989) suffices for its solution. Our formulation becomes in fact doubly easy to implement if one takes advantage of the implicit backtracking (not to mention pattern matching) built into the interpreter of the Prolog language. Our more recent implementations have therefore been in Prolog, and this has resulted in a marked improvement in program clarity (Nadel and Lin, 1991*b*).

Indications are that more complete versions of the transmission design problem are also likely to be amenable to a CSP formulation. The resulting CSP instances will however be harder to solve (more variables, values and constraints) and for these the use of Forward Checking and the like will no doubt be crucial. Prolog will then loose the advantage resulting from its implicit backtracking, although of course any of the more sophisticated approaches can be implemented in Prolog just as they can in any other language such as Lisp. Adding implicit Forward Checking say, to a Prolog interpreter would extend to harder constraint satisfaction problems the usefulness of its current implicit backtracking. Such augmented Prologs are in fact under development, as discussed in (Van Hentenryck and Dincbas, 1986) and (Van Hentenryck, 1989). For our present problem however, the Backtracking algorithm is adequate and provides the clearest demonstration of how the variables, values and constraints of our formulation work together to give transmission solutions. In Figure 9 we therefore show the action of a Backtrack search applied to our CSP formulation.

Remember, we are concerned with finding two-planetary, two-hard-link transmissions for which, by appropriate changes in input and braked part, it is possible to realize speeds $i = 0$, 1, 2 and 4 while satisfying the corresponding gear ratio ranges and other constraints of Section 4. These aspects of our transmissions can, as described in Section 3, be characterized in terms of a total of only ten CSP variables: two hard link variables $L1$ and $L2$ and, for each of the four speeds of interest, a braked-part variable Br_i and an input-part variable In_i. The output part is considered fixed at $Out = r2$ for all speeds in all transmissions. The ten CSP variables we use correspond to the ten levels of the search tree in Figure 9, with correspondences as indicated in the 'Variables' column at the left of the figure. At a given level the figure shows all possible instantiations (assignments) of that level's variable, being all the values from the domain of that variable as described in Section 3.

For extra clarity we include, in the 'Constraints' column at the left of the figure, the list of constraints (using the constraint labels of Section 4) that are checkable at each level. These are listed top to bottom in the order in which they are checked by our algorithm. To save space, some related constraints are grouped together on the same row of the figure. Under each tree node we show for each of the checkable constraints (or group of constraints), whether the test of the corresponding constraint (or group) succeeds or fails, denoted respectively by a check mark or a cross. Of course, once the first constraint-check failure occurs down the list of constraints at a node, no further constraints need be checked and the corresponding path through the tree is 'pruned off'. Only nodes at which all checkable constraints are satisfied can be used to sprout descendant nodes at the next level.

Note the six nodes which survive at level 2 (corresponding to variable $L2$) of the tree. As required, the first four are those of row 1 of Figure 8 and the second two are from row 6 of that figure. The other $18 - 6 = 12$ pairs of $L1$ and $L2$ shown in Figure 8 would also have appeared as survivors in our search tree of Figure 9 if there had been room to show all paths through to level 2.

The remaining eight levels of the tree are divided into four groups of two, corresponding to the two variables In_i and Br_i at each of the four speeds $i = 0$, 1, 2, 4. Note that checking the ratio-range constraints, C_9, C_{10}, C_{11} and C_{13}, is left till last at the corresponding Br_i levels, so as to first detect violations, if any, of the other checkable constraints. This is because checking ratio-range constraints is relatively costly. As mentioned in Section 4.2, checking a ratio-range constraint involves solving a set

https://doi.org/10.1017/S0890060400002651 Published online by Cambridge University Press

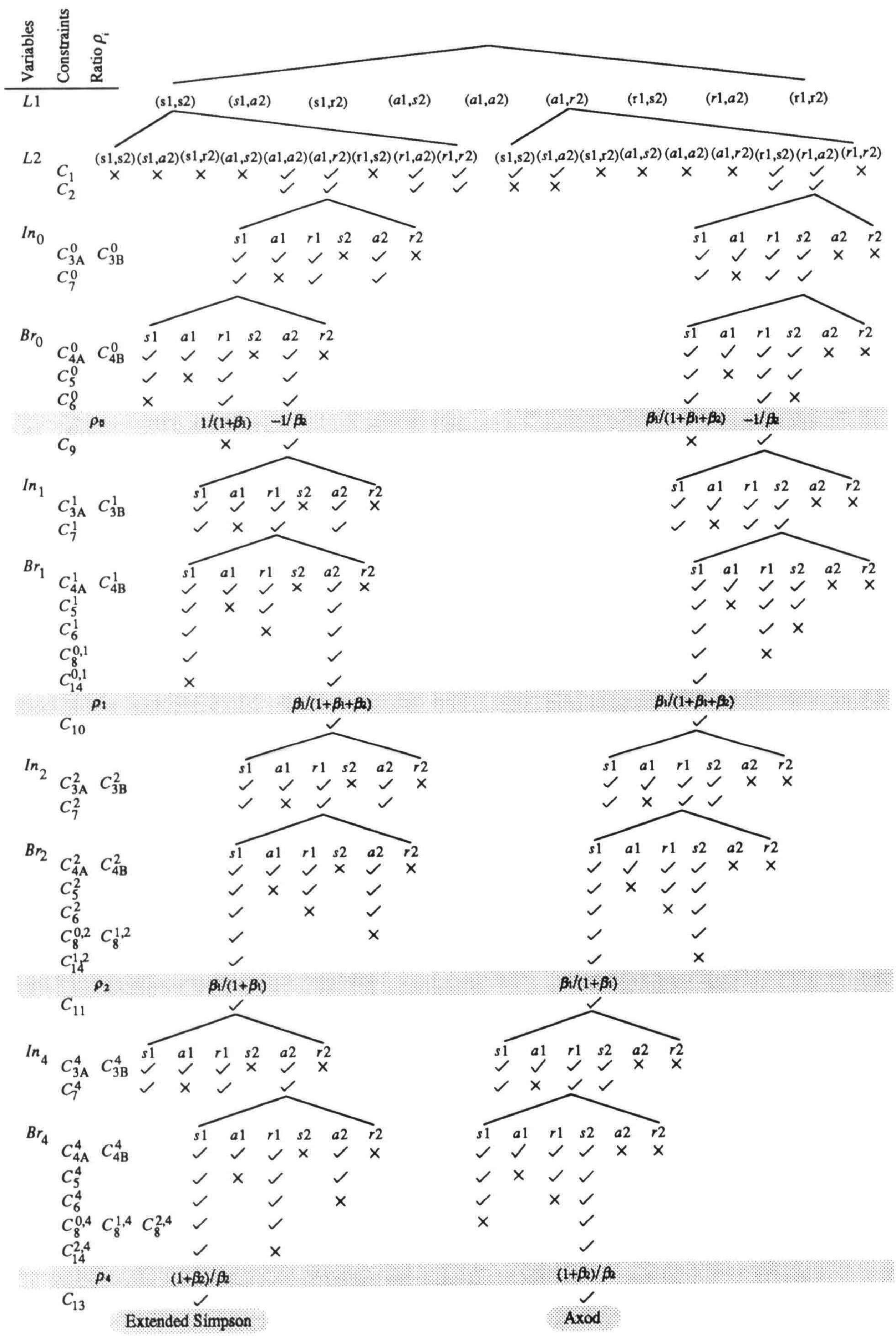


FIGURE 9. Part of the Backtrack search tree for designing 4R-speed transmissions with two planetary gearsets and two hard links.

https://doi.org/10.1017/S0890060400002651 Published online by Cambridge University Press

of simultaneous linear equations characterizing the transmission configuration, obtaining a symbolic expression for the corresponding gear ratio ρ_i and using Ward's method of equation (14) to see if there is possible overlap with the acceptable range for ρ_i. As an aid to the reader, at nodes in Figure 9 where the processing gets as far as checking a ratio-range constraint for a gear ratio ρ_i, we precede the result of that check by the symbolic expression (shown in the corresponding shaded rectangle) obtained for that ratio.

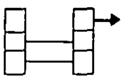
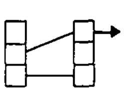
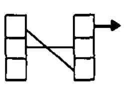
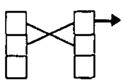
For example, there are four nodes at level 4 (variable Br_0) at which the algorithm gets as far as checking C_9, the ratio-range constraint for ρ_0. The corresponding symbolic expressions for ρ_0 are given in the first (topmost) shaded rectangle of the tree in Figure 9. At the next row of the tree we see that two of these, $\rho_0 = 1/(1+\beta_1)$ and $\rho_0 = \beta_1/(1+\beta_1+\beta_2)$, are found to violate C_9. This is obviously correct because these latter ratios are positive whereas C_9 requires ρ_0 to be negative (it is the ratio for reverse gear). The other two surviving nodes both have the same ratio $\rho_0 = -1/\beta_2$, which is found acceptable. This is also obviously correct because we can use say $\beta_2 = 3$, in the range allowed by (13), to obtain $\rho_0 = -1/\beta_2 = -1/3$, thus satisfying C_9. In general, manually confirming the ratio-range checks of our algorithm is not that straightforward. However, Table A1 of the appendix can be helpful in this regard since it gives the numerical ratio ranges, and corresponding speeds allowed by the ratio-range constraints, for all possible states generated in the search.

6.2 THE SOLUTIONS

Figure 9 shows two branches leading to solutions in our search tree. The left branch shown leads to the discovery of a 4R-speed extension of the classic 3R-speed Simpson transmission of Figure 5 and the right branch leads to discovery of the well-known 4R-speed Axod transmission of Figure 6. Interestingly, these two transmissions have the same gear ratios at corresponding speeds. The full set of solutions found by our search is shown in Table 5. Only seven solutions are found in all. Note that besides solutions 3 and 7 ('Extended Simpson' and Axod respectively) having the same set of ratios, so do solutions 2 and 6.

The number of solutions found, seven, is surprisingly small given the large size of our search space (over 136 million candidate solutions, as discussed in Section 5), and the fact that we have not yet

TABLE 5. The seven 4R-speed transmission designs (at the kinematic level) found by our program, assuming $Out = r2$, ring-to-sun teeth-ratios β_1 and β_2 bounded as given in equation (13), and ranges for the gear ratios ρ_i as given by constraints C_9, C_{10}, C_{11} and C_{13}

Solution #	$L1$	$L2$	Out	Diagram	Speed 0 (reverse) In_0	Br_0	ρ_0	Speed 1 (low underdrive) In_1	Br_1	ρ_1	Speed 2 (high underdrive) In_2	Br_2	ρ_2	Speed 4 (overdrive) In_4	Br_4	ρ_4
1 (old 1)	$(s1, s2)$	$(a1, a2)$	$r2$		$s1$	$a1$	$-\frac{1}{\beta_2}$	$r1$	$a1$	$\frac{\beta_1}{\beta_2}$	$r1$	$s1$	$\frac{\beta_1+\beta_1\beta_2}{\beta_2+\beta_1\beta_2}$	$a1$	$s1$	$\frac{1+\beta_2}{\beta_2}$
2 (old 4)	$(s1, s2)$	$(a1, r2)$	$r2$		$s1$	$a2$	$-\frac{1}{\beta_2}$	$s1$	$r1$	$\frac{1}{1+\beta_1}$	$a2$	$r1$	$\frac{1+\beta_2}{1+\beta_1+\beta_2}$	$a2$	$s1$	$\frac{1+\beta_2}{\beta_2}$
3 (old 5) Extended Simpson					$s1$	$a2$	$-\frac{1}{\beta_2}$	$r1$	$a2$	$\frac{\beta_1}{1+\beta_1+\beta_2}$	$r1$	$s1$	$\frac{\beta_1}{1+\beta_1}$	$a2$	$s1$	$\frac{1+\beta_2}{\beta_2}$
4 (old 6)	$(a1, a2)$	$(r1, s2)$	$r2$		$r1$	$a1$	$-\frac{1}{\beta_2}$	$s1$	$a1$	$\frac{1}{\beta_1\beta_2}$	$s1$	$r1$	$\frac{1+\beta_2}{\beta_2+\beta_1\beta_2}$	$a1$	$r1$	$\frac{1+\beta_2}{\beta_2}$
5 (old 8)					$r1$	$a1$	$-\frac{1}{\beta_2}$	$r1$	$s1$	$\frac{\beta_1\beta_2-1}{\beta_2+\beta_1\beta_2}$	$a1$	$s1$	$\frac{\beta_1\beta_2-1}{\beta_1\beta_2}$	$a1$	$r1$	$\frac{1+\beta_2}{\beta_2}$
6 (old 9) HM 700	$(a1, r2)$	$(r1, a2)$	$r2$		$s2$	$r1$	$-\frac{1}{\beta_2}$	$s1$	$r1$	$\frac{1}{1+\beta_1}$	$s1$	$s2$	$\frac{1+\beta_2}{1+\beta_1+\beta_2}$	$r1$	$s2$	$\frac{1+\beta_2}{\beta_2}$
7 (old 10) Axod					$s2$	$r1$	$-\frac{1}{\beta_2}$	$s2$	$s1$	$\frac{\beta_1}{1+\beta_1+\beta_2}$	$r1$	$s1$	$\frac{\beta_1}{1+\beta_1}$	$r1$	$s2$	$\frac{1+\beta_2}{\beta_2}$

same set of gear ratios ρ_i

same set of gear ratios ρ_i

Not all relevant constraints have been applied. When they are, we expect all or most of the solutions 1 to 5 to disappear, but leaving the well-known solutions 6 and 7, HydraMatic 700 and Axod. Solution numbers in parentheses are the old numbers from (Nadel and Lin, 1991*a*) and the IJCAI-91 workshop version of (Nadel and Lin, 1991*b*), where ten solutions were originally found; see footnote 2, p. 138.

incorporated quite a few relevant constraints. In fact, according to our domain experts, only solutions 6 and 7 of Table 5 correspond to commercially manufactured transmissions, the HydraMatic 700 and Axod. It is possible that some of the other five solutions we found are actually new discoveries (we haven't had time to check thoroughly), but it is more likely they will all be eliminated when the other applicable constraints are added.

For instance, the switching network (which we are ignoring here) required to build the Extended Simpson transmission is thought to not be topologically realizable. If so, we expect it to disappear as a solution when we implement topology (Section 4.5) as part of our designs. Also the symbolic, non-numeric way we are testing the ratio range constraints, based on (14), ensures only that for each speed $i = 0$, 1, 2 and 4 *independently* there are possible values for β_1 and β_2 from their allowed range (13) that allow the gear ratio ρ_i to be in its required range. Our tests do not ensure that the *same* pair of β_1 and β_2 values will work for all the ratios ρ_0 to ρ_4,[9] as is of course required in practice because the gears in our gearsets, and hence β_1 and β_2, do not change with speed for a given transmission. The latter problem will disappear once we start to explicitly consider teeth number for gears. We will then be able to test *numerically* the whole set of ratio-range constraints to ensure that their respective target ranges are satisfied for a common pair of β_1 and β_2 values. This may very well eliminate some of our current solutions 1 to 5. The step and span constraints (Section 4.2) are also most easily tested once we assign numerical values to β_1 to β_2. These may eliminate further solutions.

Besides finding transmission 'solutions' that probably don't exist, several known transmissions were in fact not found. Again, this is not a problem. It is largely the result simply of having excluded $a2$ as a possible value for the output, as mentioned in Section 3.3. This was done just as a convenient simplification, and may easily be rectified. A recent preliminary run of an extended formulation allowing also $Out = a2$ does result in the inclusion of all or most of the required remaining solutions.

[9] For example, solution 7 of (Nadel and Lin, 1991*a*), and the IJCAI-91 workshop version of (Nadel and Lin, 1991*b*) had $\rho_1 = (\beta_1\beta_2 - 1)/\beta_1\beta_2$ and $\rho_2 = (\beta_1\beta_2 - 1)/(\beta_2 + \beta_1\beta_2)$. Clearly these cannot fall in their required target ranges ($\frac{1}{3} \leq \rho_1 \leq \frac{1}{2}$ and $\frac{3}{5} \leq \rho_2 \leq \frac{4}{5}$) when using the same pair of β_1 and β_2 values in the two ratios, because the larger denominator of ρ_2 would make ρ_2 always less than ρ_1. This transmission is no longer a solution for the present formulation, but due to a different reason, as mentioned in footnote 2, p. 138.

7. Extensions

There are many ways in which our work to date may be extended. The following is a partial list, structured in terms of increasing the depth (or granularity or level of detail) of the search space, increasing the breadth (or generality) of the search space and adding an optimization capability over the search space.

(a) *Increased depth*: As mentioned in Section 6, five of the seven transmission design solutions found by our program (Table 5) are probably not physically realizable. This is because we have not yet formulated the problem at a sufficient level of detail to incorporate all relevant constraints. Even for the current level of detail all relevant constraints have not yet been incorporated. Amongst other things, our future formulations will need to consider

- teeth number on gears and the corresponding gearing constraints of Section 4.4
- the specific nature of the transmission's switching network (including clutches and brakes) and the corresponding topological constraints of Section 4.5
- the span and step constraints of Section 4.2
- the nature of the transmission's direct drive and neutral states with corresponding full use of the applicable simplicity-of-switching constraints of Section 4.3
- we mentioned in Section 2.2 that automobile automatic transmissions consist of three interacting subsystems: mechanical, hydraulic and electronic, the latter two being needed to control the former. Our present work concentrates exclusively on the mechanical subsystem. In the long run we expect to extend our transmission design task to include the integrated design of all three subsystems.

(b) *Increased breadth*: Apart from adding detail to our designs, we also aim to broaden the design space being searched, that is, generalize the class of transmissions allowed. The current search is restricted to transmissions of four forward speeds and one reverse speed, made of two simple planetary gearsets joined by two hard links and no soft (or clutchable) links. In our future formulations we expect to

- allow an arbitrary number g of gearsets, rather than just two as here
- allow an arbitrary number h of hard links, rather than just two as here
- allow an arbitrary number s of soft (or clutchable) links between gearsets, rather than none as here
- allow an arbitrary number n of (forward) speeds, rather than just four as here
- allow 'compound' planetary gearsets, rather than

https://doi.org/10.1017/S0890060400002651 Published online by Cambridge University Press

just simple planetary gearsets. A description of compound gearsets is beyond the scope of this paper. See for example Husselbee (1986) or Lynwander (1983).

(c) *Optimality*: Our present formulation of transmission design does not concern itself with optimality of the transmissions generated. At this stage we were content to confirm that we could generate a sensible feasible region (see Section 2.1), and in any case the feasible region was very small (seven solutions) so that automated optimization was not necessary. It is not clear whether the feasible region grows appreciably for the more detailed formulations we expect to use later.

Adding more variables for such features as teeth number and clutch and brake positions, will no doubt result in multiple solutions for some of the classes we are currently generating. On the other hand, some of the classes will also no doubt be eliminated due to the corresponding new constraints that will become applicable. If the net effect is that the overall number of solutions remains small in future formulations it will still not be necessary to implement automated optimization in our search algorithm. Our domain experts would be able to do it 'manually'. This would be a relief, since there are many features (such as gear noise, gear wear, cost, ease of maintenance, spread of gear ratios achieved etc.) that may be subject to optimization in designing transmissions. Extracting, formalizing and realistically balancing the relevant optimality knowledge for computer implementation would no doubt be a difficult, and relatively ill-defined, job.

If this were necessary however, we would then be solving an optimizing extension of standard CSP. We could use a single criterion function that is some weighted sum of the various features we are interested in optimizing. However since the trade-off between features is ill-defined, it would be more natural to use a *multi-criterion* optimization formulation where no *a priori* decision is made regarding relative importance of features. Navinchandra (Navinchandra and Marks, 1987*a, b*; Navinchandra, 1991) has studied such a multi-criterion optimizing extension of CSP, and in fact has done so in connection with applications, like ours, to automated design. Several other extensions of CSP, which may be relevant to our future need to optimize, have also received attention (Freuder, 1989; Rosenfeld *et al.*, 1976; Shapiro and Haralick, 1981; Ullman, 1979).

Basically the present 'first pass' at transmission design has been encouraging. We have been able to automate the rediscovery of the known transmissions within the class we have delimited, and to avoid the generation of most unacceptable solutions. The space we have used has been relatively well explored manually by human designers in the past. However, the space corresponding to the above anticipated extensions has not been manually explored to the same extent. Preliminary indications are that a CSP approach, like that used here, will successfully scale up to handle most, if not all, of the anticipated increases in breadth and depth of our formulation. The distinct possibility exists for the discovery of new and better transmissions in an automated search of such extended design spaces.

Acknowledgements

This work has been made possible by a research grant, and generous co-operation, from the Ford Motor Co. In particular I would like to thank Robert Roethler and the rest of the Ford Advanced Transmission Design Department in Livonia, Michigan, for extensive discussions on the engineering aspects of this work. Access to the Ford Motor Co. library in Dearborn was also very helpful and much appreciated.

References

Allen, J. F., 1983. Maintaining knowledge about temporal intervals *Communications ACM*, **26** (11), 832–843.

Barrow, H. G. and Tenenbaum, J. M., 1976. MSYS: A system for reasoning about scenes. *Technical Note 121, Artificial Intelligence Center,* Stanford Research Institute.

Brown, C. and Purdom, P. W. Jr., 1981. An average time analysis of backtracking. *SIAM Journal of Computers* **10,** (3), 583–593.

Cohen, P. R. and Feigenbaum, E. A. eds. 1986. *The Handbook of Artificial Intelligence,* Vol. 3. Reading, MA: Addison-Wesley.

Davis, L. S. and Rosenfeld, A. 1981. Cooperating processes for low-level vision: a survey. *Artificial Intelligence,* **17,** 245–263.

Dechter, R. 1987. A constraint-network approach to truth-maintenance. *Tech. Report R-870009, Cognitive Systems Lab., Computer Science Dept., U.C.L.A.,* Los Angeles.

Dechter, A. and Dechter, R., 1987. Minimal constraint graphs. *Proceedings of the National Conference on Artificial Intelligence (AAAI-87),* Seattle, Washington.

Dechter, R. and Pearl, J., 1988. Network-based heuristics for constraint-satisfaction problems. *Artificial Intelligence,* **34,** 1–38. Also in *Search in Artificial Intelligence,* ed. by Kanal, L. and Kumar, V. pp. 370–425, 1988, New York: Springer-Verlag.

DeKleer, J. 1986. An assumption-based TMS. *Artificial Intelligence,* **28,** 127–162.

Doyle, J. 1979. A truth maintenance system. *Artificial Intelligence,* **12,** 231–272.

Eastman, C. 1972. Preliminary report on a system for general space planning. *Communications ACM,* **15,** 76–87.

Ellinger, H. E. 1983. *Automechanics,* 3rd ed. Englewood Cliffs, NJ: Prentice Hall.

Fikes, R. E. 1970. REF-ARF: A system for solving problems stated as procedures. *Artificial Intelligence,* **1,** 27–120.

Ford Motor Co. 1985. *AXOD Automatic Overdrive Transaxle—Operation and Diagnosis.* Ford Parts and Service Division, Training and Publications Dept., order no. 1701–205.

Fowler, G., Haralick, R., Gray, F. G., Feustel, C. and Grinstead, C. 1983. Efficient graph automorphism by vertex partitioning, *Artificial Intelligence,* **21,** 245–269. Also in book: *Search and Heuristics* (1983) Amsterdam: North-Holland.

Frayman, F. and Mittal, S. 1987. Cossack: A constraints-based expert system for configuration tasks. *Knowledge-based Expert Systems in Engineering: Planning and Design,* ed. by Sriram, D. and Adey, R. A. Computational Mechanics Publications, 143–166.

Freuder, E. C. 1978. Synthesizing constraint expressions, *Communications of the ACM,* **21,** 958–966.

Freuder, E. C. 1982. A sufficient condition for backtrack-free search. *JACM,* **29** (1) 24–32.

Freuder, E. C. 1989. Partial constraint satisfaction. *Proceedings Eleventh International Joint Conference on Artificial Intelligence (IJCAI-89),* Detroit, Michigan, 278–283.

Gaschnig, J. 1974. A constraint satisfaction method for interference making. *Proceedings of the 12th Annual Allerton Conference on Circuit System Theory.* University of Illinois, 866–874.

Gaschnig, J. 1977. A general Backtracking algorithm that eliminates most redundant tests. *Proceedings of the 5th International Joint Conference on Artificial Intelligence.* Cambridge, MA.

Gaschnig, J. 1978. Experimental case studies of Backtrack vs. Waltz-type vs. new algorithms for satisficing assignment problems. *Proceedings of the 2nd Biennial Conference Canadian Society for Computational Study of Intelligence,* Toronto, Ontario.

Gaschnig, J. 1979. Performance measurement and analysis of certain search algorithms. Department of Computer Science, Carnegie-Mellon University. Ph. D. dissertation.

Hadland, T. 1987. *The Sturmey-Archer Story.* London: Pinkerton.

Haralick, R. M., Davis, L. S. and Rosenfeld, A. 1978. Reduction operations for constraint satisfaction. *Information Sciences,* **14,** 199–219.

Haralick, R. M. and Shapiro L. G. 1979. The consistent labeling problem: part I. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* **PAMI-1** (2), 173–184.

Haralick, R. M. and Elliot, G. L. 1980. Increasing tree search efficiency for constraint satisfaction problems. *Artificial Intelligence,* **14,** 263–313.

Haralick, R. M. and Shapiro, L. G. 1980. The consistent labeling problem: part II. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* **PAMI-2** (3), 193–203.

Heldt, P. M. 1955. *Torque Converters or Transmissions,* 5th ed. Philadelphia: Chilton.

Husselbee, W. L. 1986. *Automatic Transmissions: Fundamentals and Service,* 2nd ed. Englewood Cliffs, NJ: Prentice-Hall.

Jones, E. and Route, W. D. 1963. Design considerations in gear noise control. In *Noise Control in Automobile Helical Gears.* American Gear Manufacturers Association (AGMA), 299.02.

Jones, E. L. 1988. Transmission gear design for strength and surface durability. In *Design Practices—Passenger Car Automatic Transmissions,* revised 2nd ed. Advances in Engineering, Vol. 5, Society of Automotive Engineers, PA.

Juvinall, R. C. 1983. *Fundamentals of Machine Component Design.* New York: Wiley.

Kelly, O. K. 1959. The design of planetary gear trains. *Society of Automotive Engineers Annual (SAE) Meeting,* Detroit. A reduced version appears in *Design Practices—Passenger Car Automatic Transmissions,* revised 2nd ed. Advances in Engineering, Vol. 5, Soc. of Automotive Engineers, PA, 1988.

Lauriere, J. 1978. A language and a program for stating and solving combinatorial problems. *Artificial Intelligence,* **10,** 29–127.

Lynwander, P. 1983. *Gear Drive Systems: Design and Application.* New York: Marcel Dekker.

Mackworth, A. K. 1977*a*. Consistency in networks of relations. *Artificial Intelligence,* **8,** 99–118.

Mackworth, A. K. 1977*b*. On reading sketch maps. *Proc. 5th International Joint Conference on Artificial Intelligence,* pp. 598–606. Cambridge, MA.

Mackworth, A. K. and Freuder, E. C. 1985. The complexity of some polynomial network consistency algorithms for constraint satisfaction. *Artificial Intelligence,* **25,** 65–74.

Mackworth, A. K. 1987. Constraint satisfaction. In *Encyclopedia of Artificial Intelligence.* ed. Shapiro, S. Vol. 1. New York: John Wiley.

McGregor, J. J. 1979. Relational consistency algorithms and their application in finding subgraph and graph isomorphisms. *Information Sciences,* **19,** 229–250.

Mohr, R. and Henderson, T. C. 1986. Arc and path consistency revisited. *Artificial Intelligence,* **28,** 225–233.

Montanari, U. 1974. Networks of constraints: Fundamental properties and applications to picture processing. *Information Sciences,* **7,** 95–132.

Müller, H. W. 1982. *Epicyclic Drive Trains—Analysis, Synthesis and Applications.* Detroit, MI: Wayne State University Press.

Nadel, B. A. 1986. The consistent labeling problem and its algorithms: Towards exact-case complexities and theory-based heuristics, Computer Science Dept., Rutgers University, New Brunswick, N. J., Ph. D. dissertation.

Nadel, B. A. 1988. Precision complexity analysis: a case study using insertion sort. In review, available as technical report CSC-88-008, Department Computer Science, Wayne State University, Detroit, Michigan.

Nadel, B. A. 1989. Constraint satisfaction algorithms. *Computational Intelligence.* **5**(4), 188–224. A preliminary version appeared as Tree search and arc consistency in constraint satisfaction algorithms. In *Search in Artificial Intelligence,* ed. by Kanal, L. and Kumar, V. New York: Springer-Verlag, 1988.

Nadel, B. A. 1990*a*. Representation selection for constraint satisfaction: a case study using *n*-queens. *IEEE Expert,* **5**(3).

Nadel, B. A. 1990*b*. The complexity of constraint satisfaction in Prolog. *Proceedings of the Eighth National Conference on Artificial Intelligence (AAAI-90),* Boston, MA, pp. 33–39. An expanded version is available as technical report CSC-89-004, 1989, Dept. Computer Science, Wayne State University, Detroit, Michigan.

Nadel, B. A. 1990*c*. Some applications of the constraint satisfaction problem, in review. Available as technical report CSC-90-008, Dept. Computer Science, Wayne State University, Detroit, Michigan. Also in *Proceedings of the Workshop on Constraint Directed Reasoning,* and in the booklet for the tutorial *Constraint Reasoning: Theory and Applications* by Mittal, S. and Nadel, B. National Conference on Artificial Intelligence (AAAI-*90*), Boston, MA.

Nadel, B. A. 1991*a*. The complexity of backtracking and forward checking: search-order and instance specific results. In review, available from author upon request.

Nadel, B. A. 1991*b*. Automobile transmission design as a constraint satisfaction problem: modeling the gear level. A preliminary version is available from the author.

Nadel, B. A. 1991*c*. Automobile transmission design as a constraint satisfaction problem: modeling the topological level. A preliminary version is available from the author.

Nadel, B. A. and Lin, J. 1991*a*. Automobile transmission design as a constraint satisfaction problem: First results. *Proceedings 7th*

IEEE Conference on Artificial Intelligence Applications, Miami Beach, FL, pp. 248–256. Also reprinted in *Artificial Intelligence in Engineering Design,* ed. by Tong, C. and Sriram, D. Reading, MA: Addison-Wesley, 1992 (to appear).

Nadel, B. A. and Lin, J. 1991*b*. Automobile transmission design: a constraint satisfaction formulation and Prolog implementation. In review, available from authors upon request. A preliminary version appeared in *Proceedings of a Workshop on Artificial Intelligece in Design* (UCAI-*91*), Sydney, Australia, 1991, pp. 207–226.

Navinchandra, D. and Marks, D. H. 1987*a*. Layout planning as a consistent labeling optimization problem. *Fourth International Symposium on Robotics and AI in Construction,* Haifa, Israel.

Navinchandra, D. and Marks, D. H. 1987*b*. Design exploration through constraint relaxation. In *Expert Systems in Computer-Aided Design,* ed. by Gero, J. pp. 481–509. Amsterdam: Elsevier.

Navinchandra, D. 1991. *Exploration and Innovation in Design: Towards a Computational Model.* New York: Springer-Verlag.

Nudel, B. A. 1982. Consistent labeling problems and their algorithms. *Proceedings of the National Conference on Artificial Intelligence* (*AAAI-82*), Pittsburgh, PA, pp. 128–132.

Nudel, B. A. 1983*a*. Consistent-labeling problems and their algorithms: expected-complexities and theory-based heuristics. *Artificial Intelligence* **21**, 135–178. Also in *Search and Heuristics* ed. by Pearl, J. Amsterdam: North Holland, 1983.

Nudel, B. A. 1983*b*. Solving the general consistent labeling (or constraint satisfaction) problem: Two algorithms and their expected complexities. *Proceedings of the National Conference on Artificial Intelligence* (*AAAI-83*), Washington D.C., pp. 292–296.

Purdom, P. W. Jr. 1982. Evaluating search methods analytically, *Proceedings of the National Conference on Artificial Intelligence* (*AAAI-82*), Pittsburgh, PA, pp. 124–127.

Purdom, P. W. Jr. 1983. Search rearrangement backtracking and polynomial average time. *Artificial Intelligence,* **21,** 117–133.

Rit, J. 1986. Propagating temporal constraints for scheduling. *Proceedings of the Fifth National Conference on Artificial Intellıgence* (*AAAI-86*), Philadelphia, PA.

Rosenfeld, A., Hummel, R. A. and Zucker, S. W. 1976. Scene labeling by relaxation operations. *IEEE Transactions on Systems, Man and Cybernetics,* **SMC-6** (6), 420–433.

Route, W. D. 1988. Gear design for noise reduction. In *Design Practices—Passenger Car Automatic Transmissions,* Advances in Engineering, Vol. 5, revised 2nd ed, Society of Automotive Engineers, PA.

Shanahan, M. and Southwick, R. 1989. *Search, Inference and Dependencies in Artificial Intelligence.* Chichester: Ellis Horwood.

Shapiro, L. G. and Haralick, R. M. 1981. Structural descriptions and inexact matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* **PAMI-3** (5), 504–519.

Smith, B. and Kelleher, G. eds 1988. *Reason Maintenance Systems and their Applications.* Chichester: Ellis Horwood.

Stefik, M. 1981. Planning with constraints (MOLGEN: Part I). *Artificial Intelligence,* **16,** 111–140.

Tsang, P. 1987. The consistent labeling problem in temporal reasoning. *Proceedings of the Sixth National Conference on Artificial Intelligence* (*AAAI-87*), Seattle, Washington.

Ullman, J. R. 1976. An algorithm for subgraph isomorphism. *JACM,* **23,** 31–42.

Ullman, S. 1979. Relaxation and constrained optimization by local processes. *Computer Graphics and Image Processing,* **10,** 115–125.

Van Hentenryck, P. and Dincbas, M. 1986. Domains in logic programming. *Proceedings of the Fifth National Conference on Artificial Intelligence* (*AAAI-86*), Philadelphia, Pennsylvania.

Van Hentenryck, P. 1989. *Constraint Satisfaction in Logic Programming.* Cambridge, MA: MIT Press.

Waltz, D. 1975. Understanding line drawings of scenes with shadows. In *The Psychology of Computer Vision,* Winston P. H., ed. New York: McGraw-Hill.

Ward, A. C. 1988. *A Theory of Quantitative Inference for Artifact Sets, Applied to a Mechanical Design Compiler.* Department of Mechanical Engineering, MIT, Doctor of Science dissertation.

Bernard Nadel is an Assistant Professor of Computer Science at Wayne State University, Detroit. He previously taught at the University of Michigan at Ann Arbor. He received a B.Sc. Hons. in Physics and Mathematics from Monash University, Melbourne, Australia (1972), an M.Sc. in Bio-Engineering from Tel-Aviv University, Israel (1976) and a Ph.D in Computer Science from Rutgers University, NJ (1986). His research interests focus on the Constraint Satisfaction Problem and its applications in Engineering Design. Dr Nadel co-taught tutorials on Constraint Directed Reasoning at IJCAI-89, AAAI-90, CAIA-91 as well as the tutorial on Artificial Intelligence in Engineering Design at IJCAI-91.

Jiang Lin is a Ph.D. student in the Computer Science Department at Wayne State University, Detroit. She received her B.Sc. in Computer Science from Peking University, China (1986), and her M.Sc. in Computer Science from Wayne State University (1991). Her research interests are in Artificial Intelligence and its application in Engineering. She is currently working with Dr Nadel on extending the transmission design work of the current paper.

Appendix

Tables 2 and 3 above were for two particular ways of using two hard links $L1$ and $L2$, to link two planetary gearsets. For these two hard link combinations the respective tables gave the six different non-direct-drive states achievable by varying the input and braked part, when the output part is $Out = r2$. For each of these states the tables gave the symbolic gear ratio expression $\rho(\beta_1, \beta_2)$, obtained by solving the set of simultaneous linear equations corresponding to that state. This appendix, in Table A1, extends this

Table A1. Gear ratios ρ achievable using two planetary gearsets with two links $L1$, $L2$ and output $Out = r2$. Ranges for ρ assume β_1 and β_2 bounded as given in equation (13)

	 Case 1: $L1=(s1,s2)$, $L2=(a1,a2)$					 Case 2. $L1=(s1,s2)$, $L2=(a1,r2)$					 Case 3: $L1=(s1,s2)$, $L2=(r1,a2)$				
	In	*Br*	ρ	Range of ρ	possible speeds i	*In*	*Br*	ρ	Range of ρ	possible speeds i	*In*	*Br*	ρ	Range of ρ	possible speeds i
1	$s1$	$a1$	$-\frac{1}{\beta_2}$	$[-0.714,-0.286]$	0	$s1$	$r1$	$\frac{1}{1+\beta_1}$	$[0.222,0.417]$	1	$s1$	$a1$	$\frac{1+\beta_1+\beta_2}{\beta_1\beta_2}$	$[-1.939,-0.653]$	—
2	$s1$	$r1$	$\frac{\beta_2-\beta_1}{\beta_2+\beta_1\beta_2}$	$[-0.333,0.25]$	0	$s1$	$a2$	$-\frac{1}{\beta_2}$	$[-0.714,-0.286]$	0	$s1$	$r1$	$-\frac{1}{\beta_2}$	$[-0.714,-0.286]$	0
3	$a1$	$s1$	$\frac{1+\beta_2}{\beta_2}$	$[1.286,1.714]$	4	$r1$	$s1$	$\frac{\beta_1}{1+\beta_1}$	$[0.583,0.778]$	2	$a1$	$s1$	$\frac{1+\beta_1+\beta_2+\beta_1\beta_2}{\beta_1\beta_2}$	$[1.653,2.939]$	4
4	$a1$	$r1$	$\frac{\beta_2-\beta_1}{\beta_2}$	$[-1.5,0.6]$	0,1,2	$r1$	$a2$	$\frac{\beta_1}{1+\beta_1+\beta_2}$	$[0.237,0.593]$	1	$a1$	$r1$	$-\frac{1+\beta_1}{\beta_2}$	$[-3.214,-0.686]$	—
5	$r1$	$s1$	$\frac{\beta_1+\beta_1\beta_2}{\beta_2+\beta_1\beta_2}$	$[0.75,1.333]$	2,4	$a2$	$s1$	$\frac{1+\beta_2}{\beta_2}$	$[1.286,1.714]$	4	$r1$	$s1$	$\frac{1+\beta_2}{\beta_2}$	$[1.286,1.714]$	4
6	$r1$	$a1$	$\frac{\beta_1}{\beta_2}$	$[0.4,2.5]$	1,2,4	$a2$	$r1$	$\frac{1+\beta_2}{1+\beta_1+\beta_2}$	$[0.407,0.763]$	1,2	$r1$	$a1$	$\frac{1+\beta_1+\beta_2}{\beta_2}$	$[1.686,4.214]$	—

	 Case 4: $L1=(s1,s2)$, $L2=(r1,r2)$					 Case 5: $L1=(s1,a2)$, $L2=(a1,s2)$					 Case 6: $L1=(s1,a2)$, $L2=(a1,r2)$				
	In	*Br*	ρ	Range of ρ	possible speeds i	*In*	*Br*	ρ	Range of ρ	possible speeds i	*In*	*Br*	ρ	Range of ρ	possible speeds i
1	$s1$	$a1$	$-\frac{1}{\beta_1}$	$[-0.714,-0.286]$	0	$s1$	$a1$	$\frac{1+\beta_2}{\beta_2}$	$[1.286,1.714]$	4	$s1$	$r1$	$\frac{1}{1+\beta_1}$	$[0.222,0.417]$	1
2	$s1$	$a2$	$-\frac{1}{\beta_2}$	$[-0.714,-0.286]$	0	$s1$	$r1$	$\frac{\beta_1+\beta_2+\beta_1\beta_2}{\beta_2+\beta_1\beta_2}$	$[1.167,1.556]$	4	$s1$	$s2$	$\frac{1+\beta_2}{\beta_2}$	$[1.286,1.714]$	4
3	$a1$	$s1$	$\frac{1+\beta_1}{\beta_1}$	$[1.286,1.714]$	4	$a1$	$s1$	$-\frac{1}{\beta_2}$	$[-0.714,-0.286]$	0	$r1$	$s1$	$\frac{\beta_1}{1+\beta_1}$	$[0.583,0.778]$	2
4	$a1$	$a2$	$\frac{1+\beta_1}{\beta_1-\beta_2}$	$(-\infty,-1.1] \cup [2.1,\infty)$	—	$a1$	$r1$	$\frac{\beta_1+\beta_2+\beta_1\beta_2}{\beta_2}$	$[2.8,7.0]$	—	$r1$	$s2$	$\frac{\beta_1+\beta_1\beta_2}{1+\beta_1+\beta_1\beta_2}$	$[0.771,0.940]$	2
5	$a2$	$s1$	$\frac{1+\beta_2}{\beta_2}$	$[1.286,1.714]$	4	$r1$	$s1$	$-\frac{\beta_1}{\beta_2+\beta_1\beta_2}$	$[-0.556,-0.167]$	0	$s2$	$s1$	$-\frac{1}{\beta_2}$	$[-0.714,-0.286]$	0
6	$a2$	$a1$	$\frac{1+\beta_2}{\beta_2-\beta_1}$	$(-\infty,-1.1] \cup [2.1,\infty)$	—	$r1$	$a1$	$-\frac{\beta_1+\beta_1\beta_2}{\beta_2}$	$[-6.0,-1.8]$	—	$s2$	$r1$	$\frac{1}{1+\beta_1+\beta_1\beta_2}$	$[0.060,0.229]$	—

	 Case 7: $L1=(s1,a2)$, $L2=(r1,s2)$					 Case 8: $L1=(s1,a2)$, $L2=(r1,r2)$					 Case 9: $L1=(s1,r2)$, $L2=(a1,s2)$				
	In	*Br*	ρ	Range of ρ	possible speeds i	*In*	*Br*	ρ	Range of ρ	possible speeds i	*In*	*Br*	ρ	Range of ρ	possible speeds i
1	$s1$	$a1$	$\frac{1+\beta_1+\beta_1\beta_2}{\beta_1\beta_2}$	$[1.367,2.224]$	4	$s1$	$a1$	$-\frac{1}{\beta_1}$	$[-0.714,-0.286]$	0	$a1$	$r1$	$1+\beta_1$	$[2.400,4.500]$	—
2	$s1$	$r1$	$\frac{1+\beta_2}{\beta_2}$	$[1.286,1.714]$	4	$s1$	$s2$	$\frac{1+\beta_2}{\beta_2}$	$[1.286,1.714]$	4	$a1$	$a2$	$-\frac{1}{\beta_2}$	$[-0.714,-0.286]$	0
3	$a1$	$s1$	$-\frac{1+\beta_1}{\beta_1\beta_2}$	$[-1.224,-0.367]$	0	$a1$	$s1$	$\frac{1+\beta_1}{\beta_1}$	$[1.286,1.714]$	4	$r1$	$a1$	$-\beta_1$	$[-3.500,-1.400]$	—
4	$a1$	$r1$	$\frac{1+\beta_1+\beta_2+\beta_1\beta_2}{\beta_2}$	$[3.086,7.714]$	—	$a1$	$s2$	$\frac{1+\beta_1+\beta_2+\beta_1\beta_2}{\beta_1+\beta_2+\beta_1\beta_2}$	$[1.052,1.210]$	—	$r1$	$a2$	$-\frac{\beta_1}{1+\beta_2+\beta_1\beta_2}$	$[-0.479,-0.149]$	0
5	$r1$	$s1$	$-\frac{1}{\beta_2}$	$[-0.714,-0.286]$	0	$s2$	$s1$	$-\frac{1}{\beta_2}$	$[-0.714,-0.286]$	0	$a2$	$a1$	$\frac{1+\beta_2}{\beta_2}$	$[1.286,1.714]$	4
6	$r1$	$a1$	$-\frac{1+\beta_1+\beta_1\beta_2}{\beta_2}$	$[-6.714,-2.086]$	—	$s2$	$a1$	$-\frac{1}{\beta_1+\beta_2+\beta_1\beta_2}$	$[-0.210,-0.052]$	0	$a2$	$r1$	$\frac{1+\beta_1+\beta_2+\beta_1\beta_2}{1+\beta_2+\beta_1\beta_2}$	$[1.149,1.479]$	4

information to all 18 of the valid hard link combinations shown in Figure 8. For each hard link pair there are six (non-direct-drive) states, making a total of $18 \times 6 = 108$ states. The table thus contains the gear ratios for all valid states generated by our search algorithm, a partial trace of which was shown in Figure 9. As discussed in Section 5, formally each of the 108 states of Table A1 corresponds to a 4-tuple ($L1$, $L2$, In, Br) that satisfies each of the constraints C_1 to C_7 of Section 4, from the $2916 = 9^2 \times 6^2$ *a priori* possible such 4-tuples.

In addition to the symbolic gear ratio expression,

TABLE A1, *continued*

	Case 10: $L1=(s1,r2)$, $L2=(a1,a2)$					Case 11: $L1=(s1,r2)$, $L2=(r1,s2)$					Case 12: $L1=(s1,r2)$, $L2=(r1,a2)$				
	In	*Br*	ρ	Range of ρ	possible speeds i	*In*	*Br*	ρ	Range of ρ	possible speeds i	*In*	*Br*	ρ	Range of ρ	possible speeds i
1	$a1$	$r1$	$1+\beta_1$	$[2.400, 4.500]$	—	$a1$	$r1$	$1+\beta_1$	$[2.400, 4.500]$	—	$a1$	$r1$	$1+\beta_1$	$[2.400, 4.500]$	—
2	$a1$	$s2$	$\frac{1+\beta_2}{\beta_2}$	$[1.286, 1.714]$	4	$a1$	$a2$	$-\frac{1+\beta_1}{\beta_1\beta_2-1}$	$[-2.500, -0.400]$	0	$a1$	$s2$	$\frac{1+\beta_1+\beta_2+\beta_1\beta_2}{1+\beta_2+\beta_1\beta_2}$	$[1.149, 1.479]$	4
3	$r1$	$a1$	$-\beta_1$	$[-3.500, -1.400]$	—	$r1$	$a1$	$-\beta_1$	$[-3.500, -1.400]$	—	$r1$	$a1$	$-\beta_1$	$[-3.500, -1.400]$	—
4	$r1$	$s2$	$\frac{\beta_1+\beta_1\beta_2}{\beta_1\beta_2-1}$	$[1.400, 3.500]$	4	$r1$	$a2$	$-\frac{1}{\beta_2}$	$[-0.714, -0.286]$	0	$r1$	$s2$	$\frac{1+\beta_2}{\beta_2}$	$[1.286, 1.714]$	4
5	$s2$	$a1$	$-\frac{1}{\beta_2}$	$[-0.714, -0.286]$	0	$a2$	$a1$	$\frac{\beta_1+\beta_1\beta_2}{\beta_1\beta_2-1}$	$[1.400, 3.500]$	4	$s2$	$a1$	$-\frac{\beta_1}{1+\beta_2+\beta_1\beta_2}$	$[-0.479, -0.149]$	0
6	$s2$	$r1$	$-\frac{1+\beta_1}{\beta_1\beta_2-1}$	$[-2.500, -0.400]$	0	$a2$	$r1$	$\frac{1+\beta_2}{\beta_2}$	$[1.286, 1.714]$	4	$s2$	$r1$	$-\frac{1}{\beta_2}$	$[-0.714, -0.286]$	0

	Case 13: $L1=(a1,s2)$, $L2=(r1,a2)$					Case 14: $L1=(a1,s2)$, $L2=(r1,r2)$					Case 15: $L1=(a1,a2)$, $L2=(r1,s2)$				
	In	*Br*	ρ	Range of ρ	possible speeds i	*In*	*Br*	ρ	Range of ρ	possible speeds i	*In*	*Br*	ρ	Range of ρ	possible speeds i
1	$s1$	$a1$	$-\frac{1+\beta_2}{\beta_1\beta_2}$	$[-1.224, -0.367]$	0	$s1$	$a1$	$-\frac{1}{\beta_1}$	$[-0.714, -0.286]$	0	$s1$	$a1$	$\frac{1}{\beta_1\beta_2}$	$[0.082, 0.510]$	1
2	$s1$	$r1$	$-\frac{1}{\beta_2+\beta_1\beta_2}$	$[-0.298, -0.063]$	0	$s1$	$a2$	$-\frac{1}{\beta_1+\beta_2+\beta_1\beta_2}$	$[-0.210, -0.052]$	0	$s1$	$r1$	$\frac{1+\beta_2}{\beta_2+\beta_1\beta_2}$	$[0.286, 0.714]$	1,2
3	$a1$	$s1$	$\frac{1+\beta_2+\beta_1\beta_2}{\beta_1\beta_2}$	$[1.367, 2.224]$	4	$a1$	$s1$	$\frac{1+\beta_1}{\beta_1}$	$[1.286, 1.714]$	4	$a1$	$s1$	$\frac{\beta_1\beta_2-1}{\beta_1\beta_2}$	$[0.490, 0.918]$	1,2
4	$a1$	$r1$	$-\frac{1}{\beta_2}$	$[-0.714, -0.286]$	0	$a1$	$a2$	$-\frac{1}{\beta_1}$	$[-0.714, -0.286]$	0	$a1$	$r1$	$\frac{1+\beta_2}{\beta_2}$	$[1.286, 1.714]$	4
5	$r1$	$s1$	$\frac{1+\beta_2+\beta_1\beta_2}{\beta_2+\beta_1\beta_2}$	$[1.063, 1.298]$	4	$a2$	$s1$	$\frac{1+\beta_1+\beta_2+\beta_1\beta_2}{\beta_1+\beta_2+\beta_1\beta_2}$	$[1.052, 1.210]$	—	$r1$	$s1$	$\frac{\beta_1\beta_2-1}{\beta_2+\beta_1\beta_2}$	$[0.286, 0.714]$	1,2
6	$r1$	$a1$	$\frac{1+\beta_2}{\beta_2}$	$[1.286, 1.714]$	4	$a2$	$a1$	$\frac{1+\beta_2}{\beta_2}$	$[1.286, 1.714]$	4	$r1$	$a1$	$-\frac{1}{\beta_2}$	$[-0.714, -0.286]$	0

	Case 16: $L1=(a1,a2)$, $L2=(r1,r2)$					Case 17: $L1=(a1,r2)$, $L2=(r1,s2)$					Case 18: $L1=(a1,r2)$, $L2=(r1,a2)$				
	In	*Br*	ρ	Range of ρ	possible speeds i	*In*	*Br*	ρ	Range of ρ	possible speeds i	*In*	*Br*	ρ	Range of ρ	possible speeds i
1	$s1$	$a1$	$-\frac{1}{\beta_1}$	$[-0.714, -0.286]$	0	$s1$	$r1$	$\frac{1}{1+\beta_1}$	$[0.222, 0.417]$	1	$s1$	$r1$	$\frac{1}{1+\beta_1}$	$[0.222, 0.417]$	1
2	$s1$	$s2$	$\frac{1+\beta_2}{\beta_2-\beta_1}$	$(-\infty, -1.1] \cup [2.1, \infty)$	—	$s1$	$a2$	$\frac{1}{1+\beta_1+\beta_1\beta_2}$	$[0.060, 0.229]$	—	$s1$	$s2$	$\frac{1+\beta_2}{1+\beta_1+\beta_2}$	$[0.407, 0.763]$	1,2
3	$a1$	$s1$	$\frac{1+\beta_1}{\beta_1}$	$[1.286, 1.714]$	4	$r1$	$s1$	$\frac{\beta_1}{1+\beta_1}$	$[0.583, 0.778]$	2	$r1$	$s1$	$\frac{\beta_1}{1+\beta_1}$	$[0.583, 0.778]$	2
4	$a1$	$s2$	$\frac{1+\beta_2}{\beta_2}$	$[1.286, 1.714]$	4	$r1$	$a2$	$-\frac{1}{\beta_2}$	$[-0.714, -0.286]$	0	$r1$	$s2$	$\frac{1+\beta_2}{\beta_2}$	$[1.286, 1.714]$	4
5	$s2$	$s1$	$\frac{1+\beta_1}{\beta_1-\beta_2}$	$(-\infty, -1.1] \cup [2.1, \infty)$	—	$a2$	$s1$	$\frac{\beta_1+\beta_1\beta_2}{1+\beta_1+\beta_1\beta_2}$	$[0.771, 0.940]$	2	$s2$	$s1$	$\frac{\beta_1}{1+\beta_1+\beta_2}$	$[0.237, 0.593]$	1
6	$s2$	$a1$	$-\frac{1}{\beta_2}$	$[-0.714, -0.286]$	0	$a2$	$r1$	$\frac{1+\beta_2}{\beta_2}$	$[1.286, 1.714]$	4	$s2$	$r1$	$-\frac{1}{\beta_2}$	$[-0.714, -0.286]$	0

The possible speeds i assume the corresponding gear ratios ρ_i have allowed ranges as given by constraints C_9, C_{10}, C_{11} and C_{13}.

for each state Table A1 also gives the numerical range of values possible for its gear ratio given that the ring-to-sun teeth-number ratios β_1 and β_2 for the two gearsets are bounded as in (13) between 7/5 and 7/2. (Tables 2 and 3 only gave a rough version of this information by giving the particular value of the gear ratios when β_1 and β_2 had their nominal values of 2.) The gear ratio ranges, for the assumed ranges of β_1 and β_2, are obtained using Ward's method as described in Section 4.2 in connection with equation (14). In addition to numerical gear ratio ranges, Table A1 also gives for each state the zero or more

https://doi.org/10.1017/S0890060400002651 Published online by Cambridge University Press

transmission speeds, from $i = 0, 1, 2$ and 4, that the corresponding ratio range allows, given the ratio-range constraints C_9, C_{10}, C_{11} and C_{13}.

Actually, Ward's method does not apply to four of the 108 states in Table A1, because their gear ratios are not monotonic in both β_1 and β_2 as required. The states with non-monotonic gear ratios are states 4 and 6 of hard link case 4 and states 2 and 5 of hard link case 16. The gear ratio ranges given in the table for these four states have been determined graphically, without use of Ward's method. Our TRANS-FORM program however assumes all gear ratios *are* monotonic in β_1 and β_2. It thus wrongly continues to use Ward's method for the four states with non-monotonic gear ratios as well. As a result it erroneously obtains for each of them that their possible gear ratios are, by equation (14), in the range $[-1.1, \infty)$. The correct range, obtained graphically, is $(-\infty, -1.1] \cup [2.1, \infty)$, as given in Table A1. Our progam thus wrongly considers each the four states to satisfy each of the constraints C_9 to C_{13} (since the ratio range it considers possible overlaps with the coresponding required ranges) and hence to be valid for all speeds $i = 0, 1, 2$ and 4 when in fact they satisfy none of these constraints (since the actual range of possible ratios overlaps with none of the required ranges) and are valid for none of the speeds.

Luckily this unwarranted leniency on the part of our program apparently does not lead to any additional transmission designs being found acceptable. This can be seen from Table 5, in that TRANS-FORM's seven solutions shown there do not include any of the non-monotonic four states. Even though these four states are themselves (wrongly) being considered acceptable for all speeds, the other constraints of the problem are apparently sufficient to exclude transmissions where they could take part. This however is just good luck in the present version of our problem. In general we need a more reliable approach to checking the ratio range constraints. This should not be difficult, as discussed in footnote 8, p. 157.

By giving all the above-mentioned information (symbolic gear ratios ρ, numberical bounds for ρ and the corresponding allowed speeds) for all valid states, Table A1 makes explicit most of the essential intermediate results computed dynamically by our algorithm. As such, Table A1 provides a convenient way of checking the results of our algorithm. Section 6.1 alluded to this in connection with the partial trace of Figure 9. The two branches in that figure correspond to cases 2 and 18 of Table A1. The reader should check the correspondence between the two branches of Figure 9 and the entries for these two cases in Table A1.

Table A1 may also be used to quickly convince oneself that most of the other 18 valid hard link combinations cannot have any solutions, simply because they do not have at least one different state capable of providing each of the required speeds $i = 0$, 1, 2 and 4. In particular, we see from Table A1 that hard link cases 3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14 and 16 all have no states capable of providing speeds 1 or 2. That leaves only hard link cases 1, 2, 6, 15, 17 and 18. Small tree searches for cases 6 and 17, using the corresponding information provided in Table A1, show that neither of these cases allow a combination of states that provide all four speeds without violating some constraint of Section 4. Similar tree searches for cases 1, 2, 15 and 18 show that these do have solutions, and that these are exactly the solutions given in Table 5 (where of course hard link cases 1, 2, 15 and 18 are the only ones represented). Interestingly, if as we expect, solutions 6 and 7 of Table 5 ultimately prove to be the only ones really acceptable, then we will have that hard link case 18 was the only one of all 18 hard link configurations that actually supported any solutions.